

2016

Information Systems in a Changing Economy and Society: MCIS2015 Proceedings

Spyros Kokolakis

University of the Aegean, sak@aegean.gr

Maria Karyda

Dept. of Information and Communication Systems Engineering, University of the Aegean, mka@aegean.gr

Euripidis N. Loukis

University of Aegean, eloukis@aegean.gr

Yannis Charalabidis

University of Aegean, Greece, yannisx@aegean.gr

Follow this and additional works at: <http://aisel.aisnet.org/mcis2015>

Recommended Citation

Kokolakis, Spyros; Karyda, Maria; Loukis, Euripidis N.; and Charalabidis, Yannis, "Information Systems in a Changing Economy and Society: MCIS2015 Proceedings" (2016). *MCIS 2015 Proceedings*. 38.

<http://aisel.aisnet.org/mcis2015/38>

This material is brought to you by the Mediterranean Conference on Information Systems (MCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in MCIS 2015 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Spyros Kokolakis
Maria Karyda
Euripidis Loukis
Yannis Charalabidis

INFORMATION SYSTEMS IN A CHANGING ECONOMY AND SOCIETY

**Proceedings of the 9th Mediterranean Conference on
Information Systems (MCIS 2015)
Samos, Greece, October 3-5, 2015**



UNIVERSITY OF THE AEGEAN

Editors

Spyros Kokolakis
University of the Aegean
Samos
Greece

Maria Karyda
University of the Aegean
Samos
Greece

Euripidis Loukis
University of the Aegean
Samos
Greece

Yannis Charalabidis
University of the Aegean
Samos
Greece

ISBN (ebook): 978-618-82208-6-7
URL: <http://aisel.aisnet.org/mcis2015/>
University of the Aegean

Contents

MCIS 2015-Main Conference

| | |
|--|------------|
| Information Systems in a Changing Economy and Society in the Mediterranean Region | 1 |
| <i>Charalabidis Yannis, Karyda Maria, Kokolakis Spyros and Loukis Euripidis</i> | |
| Designing of Adaptive Behaviour | 7 |
| <i>Dov Te'eni</i> | |
| Exploring the Role of Business Process Management in Sustainability Initiatives.. | 17 |
| <i>Olga Levina</i> | |
| Design of a New Data Structure to Support Non-Invasive Diagnostic on Heritage Metals..... | 24 |
| <i>Antoine Rosselet, Vincent Rochat, and Cédric Gaspoz</i> | |
| Smartphone Medical App Use: A survey among Medical Students at Aristotle University of Thessaloniki | 38 |
| <i>Ioannis Chatzipavlou, Nikolaos Misirlis, and Maro Vlachopoulou</i> | |
| The determinants of Cloud Computing Adoption by Large European Firms | 50 |
| <i>Xiaolin Cheng and Ahmed Bounfour</i> | |
| Agile Drafting of IT-Outsourcing Contracts..... | 62 |
| <i>Bjarne Rerup Schlichter and Kristian Kromann, and Reumert Storgaard</i> | |
| Mobility and Security in the New Way of Working: Employee Satisfaction in a Choose Your Own Device(CYOD) Environment..... | 73 |
| <i>Arjan de Kok, Yvette Lubbers, and Remko W. Helms</i> | |
| Networking Websites Usage: All Students Multitask, Do All Manage Cognitive Load? A Comparative Study | 93 |
| <i>Sana Rouis, Faouzi Ayadi, and Rim Zarbout</i> | |
| Factors Affecting Recommendations' Acceptance in Off-Line Environment | 105 |
| <i>Sofia Gkika, Maria Skiada, and George Lekakos</i> | |
| The Value of IT: Explaining the Strategic Role of Information Systems for Fast Growing SMES | 118 |
| <i>Nikolaus Obwegeser, Ana Luiza De Araújo Burcharth, and Andrea Carugati</i> | |

Electronic Government (GOV 2.0)

- Prosperity Indicators: Foundations, Concerns, and Prospects of Usage in Policy Making-The Policy Compass Approach** 131
Ourania Markaki, Panagiotis Kokkinakos, Sotiris Koussouris, Costas Koutras, and John Psarras
- Using Social Media Monitoring for Public Policy Making - An Evaluation**..... 140
Aggeliki Androutsopoulou, Yannis Charalabidis, and Euripidis N. Loukis

Open Innovation & Knowledge Management

- Open and Responsible Innovation Concepts for Competitive Advantage**..... 154
Josip Maric, Florence Rodhain, and Yves Barlette
- Learning Phenomena of MMORPG PLAYERS: A Proposed Research Model**..... 162
Antoine Chollet, Isabelle Bourdon, and Florence Rodhain
- Understanding Knowledge Sharing in Health Care System**..... 175
Iman Ben Hamouda, Mondher Feki, Imed Boughzala, and Olfa Chourabi
- Big Data Technologies: Additional Features or Replacement for Traditional Data Management Systems?** 189
Dhouha Jemal, Rim Faiz, and Sami Mahfoudhi
- Open Innovation Strategies and Web Technologies** 197
Mokhtar Amami

IS and Organisations

- Information Systems Governance as a System of Rules: Hierarchical and Heterarchical Implementation** 220
Abderrezak Tedjini, Chantal Morley, and Eddie Soulier
- The Faces of Bureaucracy: Understanding Enterprise Social Media**..... 232
Osama Mansour and Ahmad Ghazawneh
- Integrating Open Educational Resources to Foster Serious Games and Gamification Design Principles** 246
A. I. Kokkinaki, Aresti Christoforos, and Yioula Melanthiou
- Knowledge Sharing in Computerization Projects: An Approach Based on Boundary Objects**..... 253
Mouna Ben Chouikha and Salem Ben Dhaou Dakhli
- Did the Reduction of ICT Investment due to the 2008 Economic Crisis Affect the Innovation Performance of Firms? - An Exploratory Analysis Based on Firm Data for the European Glass, Ceramics and Cement Industry** 274
Spyros Arvanitis and Euripidis N. Loukis

Cloud Computing Adoption Motivation in the European North and South..... 295
Euripidis N. Loukis and Niki Kyriakou

**A Study of Users' Perception of IT Governance During Information Technology
Adoption in Organisations** 307
Osden Jokonya

**Rationality VS Fashion In Cloud Adoption Decisions: The Case of Cloud-Enabled
Payroll Systems** 319
Ariana Polyviou, Katerina Pramatarı, and Pouloudi Nancy

Crypto-Currencies and Digital Payment Systems

**Cryptocurrency Value Formation: An Empirical Analysis Leading to a Cost of
Production Model for Valuing Bitcoin.....** 335
Adam Hayes

How Perfect Offline Wallets Can Still Leak Bitcoin Private Keys..... 354
Stephan Verbücheln

Trust in Digital Currency Enabled Transactions Model..... 363
Alex Zarifis, Xusen Cheng, Salomi Dimitriou, and Leonidas Efthymiou

**Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin
Prices** 371
*Ifigeneia Georgoula, Demitrios Pournarakis, Christos Bilanakos, Dionisios N.
Sotiropoulos, and George M. Giaglis*

Big data, data protection and privacy in a global society

Extracting Greek Elections Tweet's Characteristics 383
Katerina Kalaidopoulou, Angeliki Romanou, and George Lekakos

Evaluating Privacy Practices in Web 2.0 Services..... 394
Konstantina Vemou and Maria Karyda

Privacy and the Digital Generation Gap: Myth and Reality 404
Ruth Halperin and Yuval Dror

Digital Entrepreneurship and the Future Enterprise

RFID Enabled Visualisation of Product Flows: A Data Analytics Approach 418
Anastasia Griva, Cleopatra Bardaki, and Katerina Pramatarı

The Shape of Digital Transformation: A Systematic Literature Review 431
Emily Henriette, Mondher Feki, and Imed Boughzala

| | |
|--|------------|
| Mobile Apps for Omnichannel Retailing: Revealing the Emerging Showroom Phenomenon | 444 |
| <i>Chris Lazaris, Adam Vrechopoulos, Georgios Doukidis, and Aikaterini Fraidaki</i> | |
| Challenges to Newspapers in Digital Platform Ecosystems | 461 |
| <i>Maria Åkesson and Michel Thomsen</i> | |
| From Understanding to Use and Compete: A translational Platform for Business Transformation..... | 473 |
| <i>Gianluigi Viscusi, Christopher L. Tucci, Panagiotis Kokkinakos, Iosif Alvertis, Sotiris Koussouris, and Fenareti Lampathaki</i> | |

INFORMATION SYSTEMS IN A CHANGING ECONOMY AND SOCIETY IN THE MEDITERRANEAN REGION

Charalabidis Yannis, Karyda Maria, Kokolakis Spyros, Loukis Euripidis, University of the Aegean, {yannisx, mka, sak, eloukis}@aegean.gr

Abstract

The 9th Mediterranean Conference on Information Systems (MCIS), organized in Samos on 3-5 October 2015 by the Information Systems Laboratory (ISL) of the Department of Information and Communication Systems Engineering of the University of the Aegean, under the auspices of Association of Information Systems (AIS), had a very ambitious objective: to contribute to the creation of new knowledge about the adoption, use and impact (at various levels: individual, firm, government agency, sector, society) of the numerous and continuously increasing types of information systems (IS) in the context of the rapidly changing modern economy and society. This paper provides an overview of the papers presented at the MCIS conference, identifying the main topics and issues as expressed by more than 50 authors and contributors.

1. Introduction

Many different types of IS have been developed (e.g. both ‘internal’, oriented towards the interior of the firm, such as the ERP, and ‘extrovert’ ones, oriented towards its external environment, such as e-sales, e-procurement, e-collaboration, etc.), and many more are continuously emerging (e.g. social media, business analytics, cloud, etc.), which penetrate and transform multiple areas of our economic and social life. McKinsey in a recent report (Bughin, Chui and Manyika, 2013) conclude that ‘the combined effects of emerging Internet technologies, increased computing power, and fast, pervasive digital communications are spawning new ways to manage talent and assets as well as new thinking about organizational structures’, and finally disruptive ICT ‘will transform life, business, and the global economy’.

Furthermore, they identify the disruptive information and communication technologies (ICT) based business trends, which are expected to have a strong impact in the near future: i) Social media and technologies are gradually becoming a powerful ‘social matrix’, providing an organizational infrastructure that internal links employees, customers, and suppliers as never before; ii) Big data and business analytics are evolving as critical tools that need to be deeply embedded across functions and operations, enabling firms’ management to gain a better and deeper understanding of markets and internal operations, and finally make better business decisions; iii) Internet of things, which includes the installation of tiny sensors and actuators in a big variety of physical products, devices and equipment, an also their network connection, is expected to lead to numerous revolutionary products’ and services’ innovations; iv) The cloud delivering extensive digital processing power at low cost and in small increments, will change not only the profile and functions of firms’ ICT departments, and the whole business computing landscape in general, but also will lead to the emergence of a wide range of new

business models by shifting the economics of ‘rent versus buy’ trade-offs for companies and consumers; in general there is a trend towards ‘offering anything as a service’, leading to replacing buying and selling physical products with buying and selling services derived from physical products (so buyers, both individual and organizational ones, have the opportunity to replace big blocks of capital investment with more flexible and granular operating expenditures); v) Automation of knowledge work: ICT has initially automated physical labor and transactional tasks over the last three decades, and now advances in data analytics, low-cost computer power and machine learning are moving the automation frontier rapidly toward the knowledge workers, promising a new phase of corporate (and probably government?) productivity; vi) Exponential growth of mobile technologies penetration (both in the developed and the developing economies), such as Internet-enabled smartphones and other mobile devices, will enable vast numbers of new applications and sources of value (a good example being the big success of the mobile-payment services); vii) Fast growing bi-directional interconnection and integration of the digital and physical: initially the former ‘mirrored’ the latter, but subsequently we are increasingly seeing the opposite, with the latter being enriched and ‘augmented’ by the former, as various physical world activities, from shopping to factory work, become richer with digital information (e.g. using the mobile Internet or other types of ICT); viii) New business models of free, easy to use Internet-based personalized services (e.g. concerning information seeking, shopping, reading, watching and interacting with other people); ix) Rapid evolution of the ICT-enabled commerce (e.g. through Internet, mobile devices, etc.), which is changing from an experimental into a mainstream activity, is reducing entry barriers, opening new revenue streams to a range of new entrants, individuals and firms, supporting peer-to-peer commerce among individual consumers that replaces commercial activities traditionally carried out by firms, and also giving birth to new kinds of payment systems and monetization models; x) Government, education, and health care, initially not among the pioneers in the exploitation of ICT, have now started (as ‘late adopters’) to benefit from adopting digital technologies at the same level as many business sectors have, in order to address the social imperative ‘to do more with less’, as technological innovations can improve the quality and reach of many of the services they provide, and this can have strong positive impacts on the productivity of the business sectors.

The above ICT-based business trends are evolving in the highly complex and rapidly changing context of the modern economy and society, being shaped by this context, and at the same time influencing it significantly. In particular, there is an increasing ‘globalization’ of economic activity, with more and more firms expanding their operations and market reach beyond the borders of their ‘home’ countries, and a shift from national markets to international ones (Dess et al., 2014; Wheelen and Hunger, 2012). In the past firms could be successful by focusing only on manufacturing and selling goods and services within their national markets; however, this has today radically changed: the integrated internationalization of markets drives firms to have multi-national presence, in order to reach the economies of scale required to achieve low costs of their products and services, and therefore low prices of them, which are absolutely necessary for being competitive. The worldwide availability of the Internet in combination with transport innovations, such as containerized shipping, enable firms to have multiple manufacturing facilities and warehouses in many in many different countries and work with multiple partners all over the world in order to serve any market. In general, goods, jobs, knowledge, and capital are now able to move across borders with far greater speed and far less friction than before. This offers on one hand opportunities to access larger potential markets and a broad base of production factors such as raw materials, labor, skilled technical professionals, etc., however on the other hand increases competition dramatically, which makes it necessary for firms to improve their efficiency, innovation capacity and in general their competitiveness in order to survive.

Another important feature of the modern economy is the shift of economic activity from the western world to the eastern one. Research conducted in the LSE on this (Quah, 2011) has concluded that while in 1980 the ‘center of gravity’ of the global economy was in the mid-Atlantic (reflecting the fact

that most of the world's economic activity at that time occurred in either North America or Western Europe), by 2008, due to the rise of China and other East Asian countries, the center of gravity had moved to a location east of Helsinki and Bucharest, and it continues moving eastwards, so it is expected that by 2050 it will be located between India and China. New economic powers are emerging, such as the BRICS countries (Brazil, Russia, India and China); though initially they were centers of labor and production, they are now evolving to consumption-oriented economies as well, with increasing middle classes, and this creates tremendous opportunities for consumer products' firms all over the world. A third highly important feature of the modern economy and society are the growing inequalities among citizens in most countries, with respect to income, quality of life and access to education and health care (see Dabla-Norris et al. (2015) presenting an interesting study of the IMF on this problem), and at the same time among countries as well. We experience a good example of this trend in our neighborhood: the increasing divergence between the northern and southern European countries. The European North-South divide has been one of the most important and widely debated problems of Europe for long time (Aiginger, 2013a and 2013b; Landesmann, 2013). The countries of the European South traditionally have lower levels of economic development, productivity and performance, and also higher levels of unemployment, than the countries of the European North. Though there has been a convergence between the European North and South for some time, recently, due to the economic crisis, this trend has stopped, and on the contrary a divergence is observed (Aiginger, 2013a). The Center for Board Governance of PricewaterhausCoopers have identified five 'megatrends', which are expected to affect significantly the economy and society in the near future (PricewaterhausCoopers 2014): accelerating urbanization, climate change and resource scarcity, demographic shifts (with explosive population growth in some areas, against declines in others), shift in global economic power (from the West to the East) and technological breakthroughs (in nanotechnologies, ICT, etc.).

2. The papers

It is a critical duty of the information systems (IS) research community to investigate the adoption, use and impact of these numerous types of IS, both the established 'classical' and the continuously emerging ones, in this highly complex, challenging and rapidly changing context, in order to create useful knowledge in this domain, which will assist organizations in making better use of ICT for exploiting the opportunities and addressing the threats of this context. Judging from the topics and the quality of the papers presented in the MCIS 2015 we can conclude that it contributes significantly in this direction. In particular, 37 papers have been accepted for presentation in this Conference, after a rigorous but constructive review, from 12 countries: 5 Mediterranean (Cyprus, France, Greece, Israel, Tunisia), 4 Central and Northern European (Denmark, Germany, Sweden, Switzerland), 2 American countries (Canada and USA) and South Africa.

Most of these papers (23) are dealing with critical aspects of many of the abovementioned highly important emerging types of IS. In particular, we had a full session with four papers on the digital payment systems and the crypto-currencies, with significant impact on financial and commercial activities. Adam Hayes studies the crypto-currency value formation, by conducting an empirical analysis leading to a cost of production model for valuing the bitcoin (this work received the best paper award). Ifigenia Georgoula, Demitrios Pournarakis, Christos Bilanakos, Dionisios Sotiropoulos and George Giaglis are using time-Series in combination with sentiment analysis in order to detect the determinants of bitcoin prices, enabling a better understanding of their formation mechanisms. Alex Zarifis, Xusen Cheng, Salomi Dimitriou and Leonidas Efthymiou examine the determinants of trust in digital currency enabled transactions. Finally, Stephan Verbücheln identifies security weaknesses which can result in leakage of bitcoin private keys with quite negative consequences.

Furthermore we had five papers on social media, which gradually become a critical external and internal communication and knowledge exchange tool for both firms and government agencies. Osama Mansour and Ahmad Ghazawneh examine to what extent and how the internal use of social media for communication and knowledge exchange among firm's employees is associated with formal structures and characteristics. Konstantina Vemou and Maria Karyda present an interesting and revealing analysis of the privacy practices of Web 2.0 services, which identifies discrepancies between the stated goals of their privacy enhancing tools and the actual goals these tools accomplish. Sana Rouis, Faouzi Ayadi and Rim Zarbout investigate the effect of the use of social networks by students on their academic performance, and also the factors moderating this relationship, as well as the factors affecting the former. Katerina Kalaidopoulou, Angeliki Romanou and Georgios Lekakos analyze the Twitter activity in the Greek January 2015 elections, extract interesting characteristics of them and identify distinct users groups making quite different kinds of Twitter usage for promoting their political messages. Finally, Aggeliki Androutopoulou, Yannis Charalabidis and Euripidis Loukis evaluate the use of social media monitoring by government agencies as part of their public policy formulation processes, from two important perspectives: from the 'classical' ease of use perspective, and also from a public policy – wicked social problems perspective. This latter paper belongs to the e-government session of MCIS 2015, which included also an interesting paper by Ourania Markaki, Panagiotis Kokkinakos, Sotirios Koussouris, Costas Koutras and John Psarras on the use of open data for the calculation of a wide variety of prosperity indicators, which can be quite useful for assessing the impacts of government policies and policy measures, discussing foundations, concerns and prospects of their usage.

We also had three papers on cloud computing, a highly important trend that not only transforms business computing, but also reduces the barriers to the entry of new players in many industries. Euripidis Loukis and Niki Kyriakou investigate and compare empirically the motivations for cloud computing adoption in the European North and South, and identify important differences among them, contributing to this critical European debate. Ariana Polyviou, Katerina Pramataris and Nancy Pouloudi empirically investigate the importance of not only rationality oriented factors in firms' decisions for cloud computing adoption, but also of the existing 'IT fashion' as well, revealing the role that the latter plays in the case of cloud-enabled payroll systems adoption. Finally, Xiaolin Cheng and Ahmed Bounfour are dealing with the determinants of the organizational adoption of cloud computing by European large firms. Another two papers were dealing with the mobile technologies, which become increasingly important. Chris Lazaris, Adam Vrechopoulos, Georgios Doukidis and Aikaterini Fraidaki investigate the use of mobile applications for 'omnichannel retailing, revealing an interesting emerging showrooming phenomenon. Ioannis Chatzipavlou, Nikolaos Misirlis and Maro Vlachopoulou study empirically the use of smartphone medical applications, through a survey among medical students at the Aristotle University of Thessaloniki. Two more papers are dealing with another highly important e-learning technology: the business games. Angelika Kokkinaki, Christoforos Aresti and Yioula Melanthiou describe important aspects of the design, development and integration of open educational resources concerning the use of serious games and gamification as part of University courses, aiming to foster good practices among faculty members of Higher Educational Institutions. The research presented by Antoine Chollet, Bourdon Isabelle and Florence Rodhain aims to identify the social and the professional skills that can be developed through the participation in the massively multi-players online role-playing games (MMORPG).

Finally, we also had papers dealing with a variety of other important types of IS. Anastasia Griva, Cleopatra Bardaki and Katerina Pramataris are dealing with Radio Frequency Identification (RFID), and develop an interesting method of analyzing the immense volume of RFID data, which reflecting the behavior of products in retail stores, in order to produce information for inventory availability and inventory flows at different stages of the supply chain. Mokhtar Amami examines how web technologies

can be used for supporting four different open innovation strategies he identifies. The recommender systems are the topic of a paper by Sofia Gkika, Marianna Skiada and George Lekakos, investigating the factors that affect consumers' acceptance of the recommendations produced by these systems used in traditional stores. Iman Ben Hamouda, Mondher Feki, Imed Boughzala and Olfa Chourabi focus on the knowledge sharing systems used in the healthcare, investigating the barriers to their usage, and proposing solutions for overcoming them. The big data technologies are critically reviewed by Jemal Dhouha, Faiz Rim and Sami Mahfoudhi, and this leads to the identification of the main challenges in this area. Also, the design of a new data structure aiming to support non-invasive diagnosis on heritage metals is described in a paper by Antoine Rosselet, Vincent Rochat and Cedric Gaspoz.

From the remaining (14) papers, three discuss the association between ICT and entrepreneurship, while the other eleven papers address important issues with respect to the organizational exploitation of ICT in general. In the first group the paper by Gianluigi Viscusi, Christopher Tucci, Panagiotis Kokkinakos, Iosif Alvertis, Fenareti Lampathaki and Sotiris Koussouris, describes a translational platform, designed as part of the FutureEnterprise European project, aiming to bridge academic and industrial research with Internet-based entrepreneurship and digital business innovation. The paper by Maria Åkesson and Michel Thomsen explores the challenges that incumbent content providers, such as the traditional newspapers, face in order to adapt to and realize business opportunities in the new era of the digital platforms. Emiy Henriette, Mondher Feki and Imed Boughzala conduct a comprehensive literature review in this area, focusing on digital transformation, and concluding that the latter is more than just a technological shift, having strong impact on the business models, operational processes and end-users' experience.

Dov Te' Eni presented a highly innovative paper on the design of adaptive systems that can support complex behaviors in which users adapt dynamically to the progression of the task or to changing conditions during the interaction, proposing an approach that initially identifies dimensions on which users adapt their behavior, and then determines the corresponding design implications on how the system should adjust to fit user's adaptive behavior. Arvanitis and Loukis address a highly important problem: the effects of the recent 2008 economic crisis on firms' ICT investment, and through it on firms' innovation performance; an empirical analysis reveals factors that affect the ICT investment related behavior of firms during the crisis, and also negative impacts of the reduction of ICT investment on some kinds of ICT-enabled product innovation (pro-cyclical behavior). Given the frequent failures of ICT projects and ICT outsourcing, it is quite useful that we had two relevant papers. In particular, the problem of knowledge sharing in ICT projects is addressed in a paper by Mouna Ben Chouikha and Salem Dakhli, who analyze the obstacles to knowledge sharing in these projects in terms of knowledge boundaries, and propose an approach for this based on a typology of boundary objects. Bjarne Rerup Schlichter and Kristian Storgaard propose an approach for drafting complex ICT outsourcing contracts based on the main concepts and elements of agile software development methods.

Three papers focus on two important aspects of IS management at firm level: IS governance and IS strategy. In particular, the IS governance problem at firm level is the subject of a paper by Tedjini Abderrezak, Chantal Morley and Soulier Eddie, who define it as a set of rules allowing executives and stakeholders to determine how ICT decisions are made, and then propose two contrasting models of it, a hierarchical and a heterarchical one. The perceptions of users concerning ICT Governance during ICT adoption in organizations are explored in the paper by Osden Jokonya. The IS strategies of fast growing and successful small and medium enterprises are investigated in the paper authored by Nikolaus Obwegeser, Ana Luiza De Araújo Burcharth and Andrea Carugati, leading to the identification of three generic, archetypal IS strategies, ranging from full identification with and capitalization on IS to complete

cost-based outsourcing of all IS related services.

The performance of employees and the ICT related security in a ‘Choose Your Own Device’ (CYOD) environment, which is a continuously growing trend in firms, is examined by Arjan de Kok, Yvette Lubbers and Remko W. Helms, concluding that the former increases, while the latter does not need to increase if the proper security policies are in place. The role of business process management (BPM) for the transition towards an environmentally sustainable enterprise is explored in the paper by Olga Levina, reaching the conclusion that BPM can be used not only to achieve cost benefits through more efficient processes, but also to promote and achieve an efficient use of resources as well, leading to more environmentally sustainable business operations. Ruth Halperin and Yuval Dror examine to what extent there are significant differences among younger and older Internet users as to their privacy perceptions. Finally, Josip Maric, Florence Rodhain and Yves Barlette analyze two evolving innovation related concepts, the ‘responsible innovation’ and the ‘open innovation’, aiming at a better understanding of the similarities and differences between them, and also at exploring the possibility of a fusion between these two concepts.

3. Conclusions and further steps

Taking into account the thematic richness and the quality of the papers of the MCIS 2015 we believe that it has contributed to the increase of the existing body of knowledge concerning the adoption, use and impact of the various types of IS, with a strong emphasis on the emerging novel ones, and this is quite beneficial both for practice (towards a better and wider use of ICT in private and public sector organizations, in order to meet the challenges, address the threats and exploit the opportunities of the current highly complex, demanding and rapidly changing economic and social context), and for future research in this domain.

4. References

- Aiginger, K. (2013a). ‘A New Strategy for the European Periphery’, Austrian Institute of Economic Research. WIFO Working Papers - No. 443, Vienna, Austria.
- Aiginger, K. (2013b). ‘A Southern Europe strategy based on vision and industrial policy’, *Ekonomiaz*, 82 (1), 129-169.
- Bughin, J., Chui, M., Manyika, J. (2013). ‘Ten IT-enabled business trends for the decade ahead’. McKinsey Quarterly, May 2013.
- Dabla-Norris, E., Kochhar, K., Ricka, F., Suphaphiphat, N., Tsounta, E. (2015). ‘Causes and Consequences of Income Inequality: A Global Perspective’, International Monetary Fund (IMF), USA.
- Dess, G., Lumpkin, G. T., Eisner, A. B., McNamara, G. (2014). ‘Strategic Management – Text and Cases – 7th Edition’, McGraw-Hill Education, USA.
- Quah, D. (2011). ‘The Global Economy’s Shifting Centre of Gravity’, *Global Policy*, 2(1), pp. 3-9.
- Landesmann, M.A. (2013). ‘The new North-South Divide in Europe – can the European convergence model be resuscitated?’, *WIIW Monthly Report 2013* (1), 3-13.
- PriceWaterhausCoopers (2014). ‘Five megatrends and possible implications’, Center for Board Governance.
- Wheelen, T. L., Hunger, J. D. (2012). ‘Strategic Management and Business Policy - Toward Global Sustainability – 13th Edition’. Pearson Education, USA.

DESIGNING FOR ADAPTIVE BEHAVIOUR

Complete Research

Te'eni, Dov, Tel Aviv University, Israel, teeni@tau.ac.il

Abstract

We need, increasingly, to design systems that can support complex behaviours in which users *adapt* dynamically to the progression of the task or to changing conditions during the interaction. Unfortunately, we lack procedures and methods to build systems for adaptive behaviour. This article proposes an approach that identifies dimensions on which users adapt their behaviour and then determines the corresponding design implications on how the system should adjust to fit the user's adaptive behaviour.

Keywords: Adaptive behaviour, Levels of abstractions, Human-computer interaction, Design.

1 Introduction

Increasingly, we need to design the human-computer interface of systems that support complex tasks in which users *adapt* dynamically to the progression of the task or to changing conditions during the interaction. There is abundant research on adaptive systems that adapt to the user's characteristics and preferences (Billsus et al., 2002) or to the conditions at the time of use. There is also research on how users adapt their use of systems, e.g., using more features of the system (Sun, 2012). Unfortunately, we lack procedures and methods to build systems that support the user's adaptive behaviour whilst performing a task. This article proposes an approach that identifies dimensions on which users adapt their behaviour and then determines the corresponding design implications on how the system should adjust to fit the user's adaptive behaviour. Future work will examine technical solutions based on advanced technologies for data acquisition and analytics but for now, in this first step, we demonstrate our approach with extant technologies.

A reader of a professional article, such as this article, may read it in at least two different modes. The reader can scan the article by looking at prominent pieces of information such as the titles, the first sentence of a paragraph, anything in bold or italics, tables and graphs. A second mode of reading would be to read serially consecutive pages, attempting to read sequentially large parts of the text. Most readers, start, out of habit, reading an online article (or an information webpage) in the first mode. An adaptive reader seeking, say, a practical question about designing an interface may toggle between the modes, depending on the relevance of the content to the question in mind. Moving from one mode to another is not always obvious or simple and is often triggered by external situations, but once in a certain mode, the reader reads, out of habit, in a routine that is quite automatic. This simple and common example demonstrates a dimension of reading specificity (high-level reading versus detailed-level reading) on which the reader moves up and down the levels according to considerations of speed and convenience and the likelihood of success in finding an answer.

Assuming that such moves between reading modes are advantageous, there are at least two broad design implications. One implication is the need to design a system that supports an easy transition between reading modes. A second implication is to fit the system to best serve each reading mode in its turn. For instance, when scanning the article in the high-level reading mode, the reader essentially ignores non-prominent pieces of information according to a principle called visual hierarchy (Djamasbi et al., 2011), letting them almost disappear into the background. A commonly used technique to emphasize parts of the text and downplay surrounding texts is to present the page in a fisheye view.

A third, more complex design implication of the two reading modes is to prompt the user to move effectively between modes. To do so, the system would have to rely on knowledge of the relative effectiveness of one mode versus the other, given a particular task such as finding in the article a practical design solution. For instance, knowing the reader is looking for a practical solution, the system could suggest a move to detailed-level reading when the reader gets closer to the appropriate section in the article (like a car dashboard suggesting a change of gear).

In practice, not enough is being done to design for adaptive behaviour, in part perhaps, because we need to know why and how people adapt when performing specific tasks. Furthermore, adaptive behaviour relies on complex feedback. Most interactive systems provide simple feedback to users immediately upon the user's action such as 'Input received'. In contrast, this paper talks about longer cycles of feedback such as feedback throughout a series of steps needed to complete a flight reservation. This requires looking at cognitive, affective and social aspects of thinking and communicating. We refer here to adaptive behaviour that may include, for instance, moving up and down levels of abstraction, changing moods in emotional activities, and oscillating between levels of networking. How do you support effectively such adaptive behaviour?

We propose a procedure of five steps to design systems that support adaptive behaviour:

1. Identify dimensions for adaptive behaviour in a given activity
2. Determine the consequences of transitions between states on a dimension
3. Design systems to support transitions between states
4. Fit the system to the state to which the user moved
5. Design systems that guide advantageous behaviour

The remainder of the paper develops these ideas systematically. The next section reviews the idea of fitting, statically, the human-computer interface to characteristics of the user and the task. A subsequent section, extending the static view of fit to a dynamic view, develops the above five-step procedure. It demonstrates the approach by examining two distinct cases of dimensions of change. One has to do with individuals adapting throughout a modelling session and the other colleagues adapting their patterns of interaction throughout a project. The final sections talk about extensions, implications and current developments.

2 Fit in human-computer interaction

There is substantial and expanding research on fitting the interface according to the user's characteristics or preferences, according to the purpose of using the system and according to the context in which it is used. Figure 1 shows a framework for studying 'fit' in goal-oriented human-computer interaction in which the user's goal is to accomplish a task (Te'eni, 2006). The framework takes the perspective of an individual user. In today's world of collaboration and social computing, this approach should be expanded to collective human-computer interaction too. The three main components of the framework are user, task and computer, and the fit among them can be operationalized as three human aspects,

namely physical, cognitive and affective. In a collaborative context, social fit should also be considered. Extant research has shown, in particular, that the interaction of cognition with task and computer depends on whether behaviour is habitual or mindful. This section looks primarily at habitual tasks, such as reading or following some pre-defined rule such as normalizing a data table, but not at creative tasks, such as generating a new product. A good fit among the three components is expected to lead to better performance and well-being for the user. Intuitively, a good fit is one that minimizes the effort required to perform the task, be it physical effort, such as hand movement, or cognitive effort, such as the need to translate a graph to numbers or vice versa (see for example the work on cognitive fit by Vessey and Galletta, 1991).

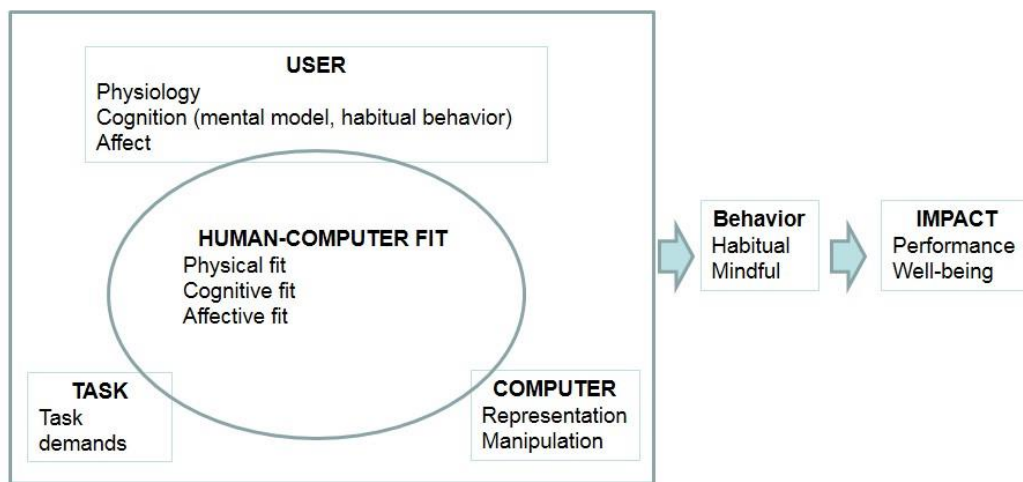


Figure 1. A framework for studying fit for an individual user (adapted from Te'eni, 2006).

Figure 1 suggests several types of human-computer fit according to the combinations of factors in a given context. For example, designers attempt to fit the screen brightness to the current conditions of light intensity (physical environment) and fit the fonts to the user's age and taste (the user's sensory capacity but also the user's preferences). These examples of fit are straightforward and rely on common knowledge and experience. However, less obvious instances of fit will have to rely on knowledge about what constitutes a good fit under particular circumstances. For instance, fitting the right format of a quantitative graphic will depend on knowledge of the task, of the graphics and of the user's cognitive capacity (Kennedy et al., 1998). Other more complex examples include 1) fitting the level of detail in a message, which would depend on the communication complexity experienced by the user, and 2) fitting the tone and facial expressions of avatars to the user's mood, which would depend on psychological knowledge of emotions. Clearly, more research is needed to determine how to adapt the interface for almost unlimited combinations of users, tasks and contexts.

The next section concentrates on one aspect of adaptation, namely adaptive behaviour during a session of human-computer interaction that involves moving between different modes of behaviour. The move between modes calls for dynamic fit rather than the static fit described above, i.e., at any point of time, it is necessary to fit the interface to the current mode of behaviour. As behaviour in each mode is assumed to be routine or well defined according to known rules, good fit will usually imply minimal interference with the expected behaviour. In contrast, the moves between modes require more complex support as shown below.

3 Dimensions of adaptive behaviour for dynamic fit

Each of the fit examples above may be regarded as a one-time adaptation, say, at the beginning of a work session with the computer. Alternatively, the system could be readapted, once the conditions change, even amidst a session. A simple example is fitting the screen brightness to the prevailing light intensity; the screen brightness can readjust automatically when the light intensity changes. Adaptation during a session is more complicated when it involves changes in the user's behaviour. For instance, when the user communicating online with a friend needs to adapt the communication during a session in order to avoid misunderstandings, it is not clear how to adapt. The user, for instance, could increase the level of message detail and provide fuller explanations, which would likely reduce errors and misunderstanding but not if the problem is the language. To understand when a user should increase message details and how the online system should be designed accordingly, we need a systematic procedure for supporting adaptive behaviour that builds on knowledge of communication.

The five step procedure mentioned in the introduction is now developed using two examples. The first example of adaptive behaviour is taken from data modelling.

1. Identify dimensions for adaptive behaviour

The first step is to determine the user's adaptive behaviour on one or more dimensions. Behaviour adapts during the activity of performing the task and therefore the activity delineates the time-related scope of adaptation. In the case of data modelling, we assume, for simplicity, that the activity can be completed in one human-computer session (stretching the activity over several sessions complicates somewhat the design).

When building, say, a class diagram of a system, data modellers adapt their modelling behaviour on at least two dimensions. One dimension is the levels of abstraction at which the data modellers think, and the second dimension is the type of activity in which they engage, e.g., planning, scoping and testing (Srinivasan and Te'eni, 1995). Figure 2 shows the transitions of a subject on both dimensions (between levels of abstraction 0-4 and between activities A-D). The figure was established based on a protocol of a think-aloud session to solve a data modelling exercise. In combination with a measure of performance in the data modelling exercise, the pattern of moves on the two dimensions represents the knowledge base for designing for adaptive behaviour in modelling. We will concentrate on levels of abstraction only, which is highly applicable to most forms of human-computer interaction. Concretely, in data modelling, one could think of entities as one level of abstraction, properties of the entity as a lower level and clusters of entities (e.g., a vehicle is a cluster of car and truck) as a higher level. Adaptive behaviour is working on one level, say entities, and climbing to a higher level of clusters when progress loses direction or dropping to a lower level of properties when modelling or testing fails.

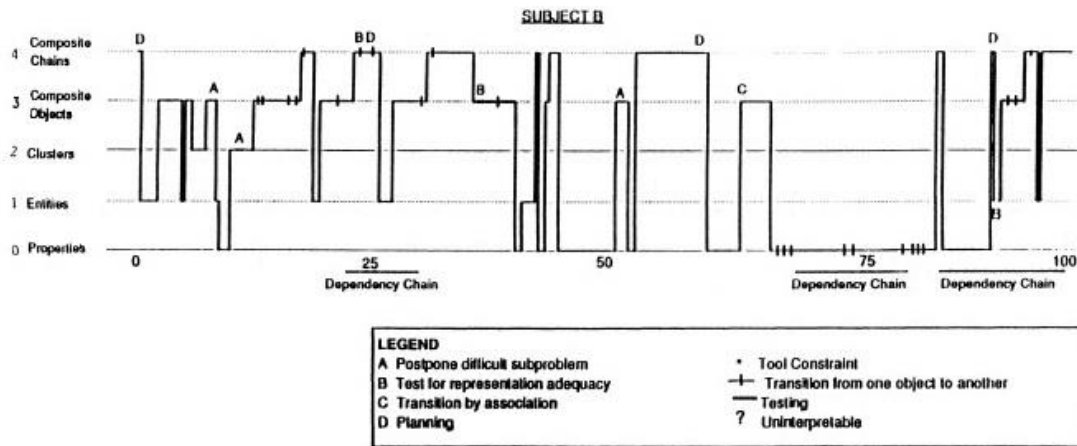


Figure 2: Adaptive behaviour when data modeling (from Srinivasan and Te'eni, 1995).

2. Determine the consequences of transitions between states on a dimension

Having identified the dimension of transitions that characterizes user behaviour, we ask what, if at all, are the consequences of certain patterns of transitions. Knowledge of data modelling suggests that certain patterns are more effective than others in problem solving. For instance, working mostly at lower levels of abstraction was less effective than working at higher levels, although working only at higher levels of abstraction was ineffective. An effective pattern was to work at the level of entities but from time to time climb to the higher levels for short periods. There are other effective patterns, which may depend on the user's level of expertise (Srinivasan and Te'eni, 1995). In any event, in this example and on the basis of knowledge about the specific phenomenon of data modelling, it is possible to determine the relationship between patterns of transitions and consequences. In other cases, we may be able to define a dimension of adaptation but we may not be able to determine the consequences of alternative patterns.

3. Design systems to support transitions between states

The most obvious design implication, and yet one that more often than not is ignored in practice, is to enable easy transition between levels of the dimension identified. In the case of levels of abstraction, this would mean, for example, an easy transition between entities and clusters of entities. If we know that users tend to move from one level to another on a dimension, the system should be design to support the transitions even if the consequences of transitions is not known (at least if it is not known to be detrimental). Easy transitions between levels of abstraction means both easily operating the system to reveal and move the focus of attention to another level and supporting cognitively the transition (Sun, 2012). The move from one level to another usually means breaking away from the habitual behaviour into which the user has settled in and requiring the user actively think how to proceed, and this requires some new and forceful condition or some external trigger (Louis and Sutton, 1991).

Technically effecting the transition to another level would usually involve some form of direct manipulation such as a single click to zoom in or an option to hover over an entity, reveal its properties and move to one of them. Cognitively supporting the transition between levels includes at least two types of support: underscoring the new state as the current focus of attention and maintaining the source as context when arriving at the target. In other words, the user often needs first to realize that the focus has shifted to the new level of properties and at the same time to see the level of entity as the context for working on the properties. For example, in Figure 3, the user has moved from the working on enti-

ties to focusing on the properties of a particular entity labeled 'Person.ContactType'. The small map of entities (the higher level of abstraction from which the user moved to the current focus) is left on the screen to present the higher level as the context for the current focus. At any time, the user can go back to the higher level by clicking on the small map.

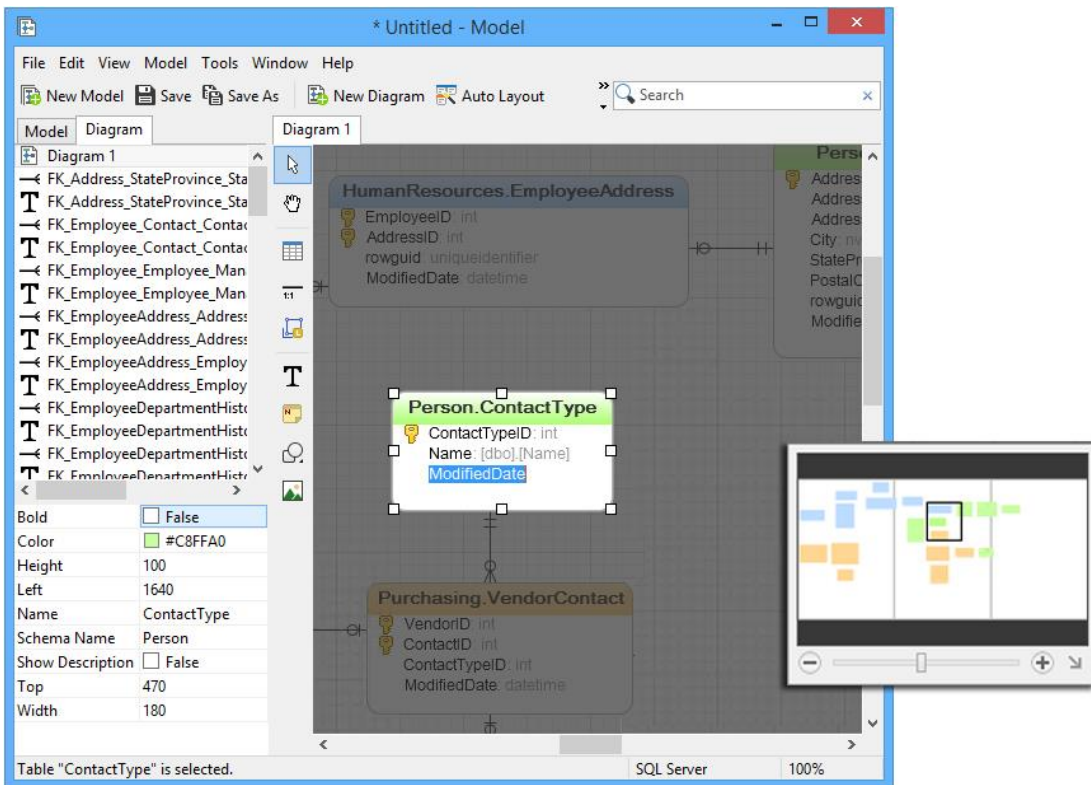


Figure 3: A screen for working at the properties level of abstraction.

4. Fit the system to the state to which the user moved

The rectangle describing the entity Person.ContactType fits the task the user is expected to perform, namely determine and check the properties of the entity, part of which are detailed in the lower left corner (e.g. Bold is set to False). A better fit would be to move the table of properties from the lower left corner closer to the central rectangle so that any moves between two require minimal effort. In fact, a clever design might be able to integrate the two in order to bring the manipulations on the entity's properties are direct as possible.

Designing for habitual behaviour differs from designing for novel behaviour that requires mindful planning and controlling. When fitting the human-computer interface to the user's expected behaviour, the designer's goal is to minimize the user's effort in performing the task correctly. When fitting the human-computer interface in novel situations, the designer's goal, in addition to minimal effort and accuracy, is to support generating new behaviours and controlling unpracticed behaviours. In the case of setting or manipulating an object's properties, behaviour is assumed to be habitual (unless the user is in training). The screen is designed to allow easy manipulations of data, immediate feedback (e.g., setting the color immediately shows the color to be displayed during use under the anticipated conditions of use), and easy and therefore more accurate detection of properties and their values.

When working at the higher level of abstraction, namely at the level of entities, the user is expected to plan the relationships between entities. The small map of entities, in an expanded form, would become the centrepiece of the screen. Its graphic presentation is a better (cognitive) fit than say a table of relationships because it lends itself most immediately to the way people represent relationships in their mind (Vessey and Galletta, 1991).

5. Design systems that guide advantageous behaviour

The last of the five steps in the procedure for designing adaptive behaviour systems requires knowledge of performance measures in the particular domain of the user's work. Furthermore, this step is feasible only if there is a convincing argument for affecting consequences advantageously by manipulating the human-computer interaction (step #2 above). The research quoted above argues that certain patterns of transitions are more effective than others are. For instance, remaining too long at the lowest levels of abstraction without occasionally taking a more comprehensive view by climbing to a higher level of abstraction will lead to errors. The system can detect relatively long periods of working on properties by monitoring the user's manual inputs to the system (in other cases, the system could monitor the user's gaze). The system can then alert, suggest, or force corrective action such as moving to a higher level. In other cases, it may be important to consider not only the proportionate time spent at different levels but also the sequence of activities.

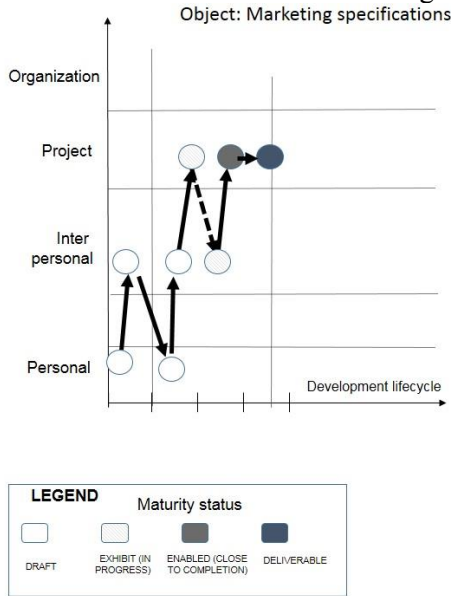
Technologies that monitor users' behaviour and conditions are becoming ubiquitous in certain domains such as health and work safety but will most likely spread to many other domains of life. The nearly constant accessibility to devices such as the cell phone, online watch and wearable devices in general coupled with knowledge on what is effective in which conditions makes it feasible to guide advantageous behaviour. A simple example is RunKeeper's (an App to plan and monitor jogging activities) real time health graph, which could easily be supplemented with alerts on when you should accelerate to meet your running goal or slow down to maintain your health. Similarly, a modelling system could, in addition to signalling when the user should move to another level of abstraction, motivate the user with graphical depictions of successful patterns vis. a vis. her own actual pattern.

4 An organizational dimension of adaptive behaviour

The second example of using the procedure demonstrates a very different situation in which the dimension of adaptive behaviour is in the pattern of collaboration. The context is innovation where project members collaborate in order to develop a new product. The time dimension is not during a session but rather during stages of a project. Nevertheless, the same principles of designing for adaptive behaviour apply as we can see in the five steps of the procedure for designing adaptive behaviour systems.

Dimension of transitions. In studying how team members work on developing new products, research has shown that people work individually, in small and trusted inter-personal groups, in formal project teams and in the set of all organizational actors that are relevant to the project. Different actors at different stages of product development in different projects will show different patterns of moving from one mode to another. Figure 4 shows a particular pattern of transitions between modes when developing the marketing specifications of a new product (Merimond et al., 2012). The reasons for moving from one mode to another vary. For instance, when innovators feel they need feedback, they move from personal to inter-personal thinking but they may be reluctant to share with the entire project team for lack of trust. Innovators seem to change their patterns of behaviour depending on the organizational climate of safety. When people feel safe, they tend to move faster to wider circles of thinking but

tend to hold back on moving to the project level when they feel un-



safe.

Figure 4: Innovating at different thinking and communicating (adapted from Merimond et al., 2012).

Consequences of patterns of transitions. Certain patterns of transitions between modes result in poor innovation or in delayed innovation. For example, when development continues at lower levels without accessing the project level early in the process, some perspectives may be ignored resulting in incomplete designs that later need to be corrected. In contrast in situations of high technological novelty, insufficient oscillations between personal and interpersonal modes results in premature formal discussions. In Figure 4, the maturity status of an object at any moment show how complete and accurate is the marketing specification. A relatively small number of oscillations between personal and interpersonal modes before the maturity status turns to the mature status ('enabled' or 'deliverable') is likely to result in poorer quality plans or products.

Design systems to support transitions. Google+ used circles to allow easy sharing of information between different groups of people regardless of what media was used. For such a system to work it, the transitions between working alone to working with a small group and between working with a small group to working with a bigger group must be easy for the user. A user who works on some object, like a marketing specification, should not be made to copy and edit it before sharing it. The system should enable a simple designation of what and with whom to share any object at any stage as a basis for communication and collaboration. Nevertheless, in some cases, the user shares information, tailoring the message to the receiving party and annotating the document with specific requests for feedback. At the same time, the system should let the user bring the discussion back to smaller circles for further 'closed' deliberations using new versions of the marketing specification.

Fit the system to the state to which the user moved. The human-computer interface for the personal mode of thinking would differ that of the interpersonal mode. In the former, the view of the marketing specification document is meant to encourage creative and critical thinking of a single user. In the interpersonal view, while the marketing specification must remain the focal object on the screen, it is necessary to add the perspectives of the interpersonal circle. For instance, it will first be important to specify who will receive the information. An example of good fit would be to show the members of the interpersonal group with pictures and names. This is the format users usually represent their close colleagues. More sophisticated interfaces help the user tailor the message to recipients, such as highlighting certain aspects that are relevant only to the particular recipient.

Design systems that guide advantageous behaviour. In this case it was difficult to propose interface designs that prompt effective patterns or discourage ineffective patterns. While it was straightforward to design easy transitions governed by the user, triggering transitions seemed risky. Our conclusion was to alert the user in two situations: when there is only one conversation at the interpersonal mode when the maturity level increase to 'close to completion' and when there are no oscillations back to individual and interpersonal modes from project level before going to the final level of maturity.

5 Beyond cognitive support

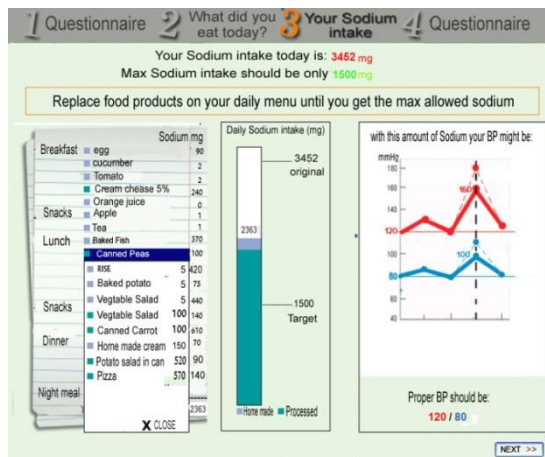


Figure 5: Personalized screen on the basis of user information (from Ronen and Te'eni, 2013).

A revealing example of the impact of personalization (adapting to user information) is in a recent study of health systems. The system supports self-management by providing ongoing feedback showing the effect of food consumption on health. The study compared personalized versus generalized information with respect to the users' attitudes and behaviour. Subjects receiving the personalized feedback (Figure 5) reported a more positive attitude and a greater propensity to follow the system's recommendations in order to improve their well-being. These findings suggest that fit is not only a means for improving the human-computer interaction by minimizing the effort to act out intended behaviour but also serves to change attitudes and behaviour.

In future, adaptive behaviour may be supported so as to achieve affective fit as well as cognitive and physiological fit. However, this would require monitoring emotions, which is not an easy task.

6 Conclusion and technological feasibility

The technological feasibility of supporting adaptive behaviour in human-computer interaction is increasing dramatically, and so will the expectations. As demonstrated above, in order to proactively recommend or even trigger adaptive behaviour, the system must recognize the user's current position on dimensions of change and the current conditions. More and more sophisticated technologies (e.g., the Internet of Things) will be available to detect information about the user, the task and the setting. For instance, a user engaged in collaborating with remote others could be notified on the susceptibility of miscommunication at the receiver's end on the basis of sensors detecting noise in the communication. The system can then recommend to increase the level of explanations. More sensors transmit more information to users about the environmental conditions as well as the bodily conditions; most probably people will not ignore the information but rather feel pressurized to adapt. The new information technologies and infrastructures necessitate but also enable more adaptation.

The five-step procedure for designing systems that support adaptive behaviour relies on knowledge of the user, task and setting. The first step is to use the knowledge of the user's behaviour when solving the task in order to determine the dimensions of changing behaviour. More knowledge and more sophisticated computing and data analytics will allow building better support for adaptive behaviour.

References

- Billsus, D., Brunk, C. A., Evans, C., Gladish, B., & Pazzani, M. (2002). Adaptive interfaces for ubiquitous web access. *Communications of the ACM*, 45(5), 34-38.
- Djamasbi, S., Siegel, M., & Tullis, T. (2011). Visual hierarchy and viewing behavior: An eye tracking study. In *Human-Computer Interaction. Design and Development Approaches* (pp. 331-340). Springer Berlin Heidelberg.
- Kennedy, M.; Te'eni, D.; and Treleavan, J. (1998). Impacts of decision task, data and display on strategies forextracting information. *International Journal of Human Computer Studies*, 48, 159–180.
- Louis, M. R., & Sutton, R. I. (1991). Switching cognitive gears: From habits of mind to active thinking. *Human relations*, 44(1), 55-76.
- Merminod, V., Rowe F., Te'eni D (2012) Knowledge Sharing and Maturation in Circles of Trust: The case of New Product Development ICIS2012, Orlando, FL, December.
- Ronen H. and Te'eni D (2013). The Impact of HCI Design on Health Behavior: The Case for Visual, Interactive, Personalized (VIP) Feedback. ICIS2013, Milan, December, 16-18, 2013
- Srinivasan A & Te'eni D (1995). Modeling as constrained problem solving: An empirical study of the data modeling process. *Management Science*, 41(3), 419-434.
- Te'eni, D. and Sani-Kuperberg, Z (2005). Levels of abstraction in designs of human-computer interaction: The case of e-mail. *Computers in Human Behavior*, 21, 817-830.
- Vessey, I., & Galletta, D. (1991). Cognitive fit: An empirical study of information acquisition. *Information systems research*, 2(1), 63-84.

EXPLORING THE ROLE OF BUSINESS PROCESS MANAGEMENT IN SUSTAINABILITY INITIATIVES

Research in progress

Levina, Olga, Technische Universität Berlin, Berlin, Germany, Olga.levina.1@tu-berlin.de

Abstract

The sustainability of business activities is gaining increasing importance with organizations putting more effort into this topic. Approaches from the field of Business Process Management (BPM) can support sustainability initiatives. This paper investigates the question raised by BPM researchers about the role of BPM for the transition towards a sustainable enterprise. Online content analysis of 78 case studies is used to derive the context, methods and effects of environmental initiatives across industries. While BPM techniques are foremost implemented and used to achieve cost benefits through more efficient processes, the results of this explorative study indicate they are also applied to promote and achieve an efficient use of resources leading to more sustainable business operations if supported by the top management and a “green” strategy. These results are industry independent and support the research findings in the domains of quality management as well as strategy realization providing BPM researchers and practitioners with insights on the assessment of a transition towards a more sustainable business.

Keywords: Business process management, sustainability, green BPM, exploratory study.

1 Introduction

Sustainability has been a topic on business agendas for several decades, from the time when the enterprises realize their impact on and responsibility towards the environment. This “going green” movement has become popular since the companies have realized that they can reduce pollution and increase profits simultaneously. The subject of sustainability is also being increasingly discussed in Information Systems Research (ISR), with green IT and green Business Process Management (BPM) being the topics that gain more and more popularity within the discipline, e.g. (Pernici et al., 2012; Molla, Cooper, and Pittayachawan, 2009; vom Brocke, Seidel, and Recker, 2012; Loos et al., 2011; Nowak et al., 2011; Loeser, 2013; Reiter, Fettke, and Loos, 2014a). BPM experts develop tools and metrics that are aimed at integrating sustainability aspects into the process analysis and (re-) design (Hoesch-Klohe and Ghose, 2012; Wesumperuma et al., 2011; Opitz et al., 2012; Recker, Rosemann, and Gohar, 2011; Nowak et al., 2011) following the conviction that “making processes more sustainable will ultimately lead to making organisations more sustainable” (Pernici et al., 2012).

This paper explores the questions of what managerial methods are used by enterprises to realize sustainability, also known as “green”, initiatives and what role BPM plays in their realization. This paper adopts the following definition of a sustainable business: a sustainable business should “[. . .] meet the needs of the present without compromising the ability of future generations to meet their own needs” (Brundtland Commission, 1987). Environmental sustainability is often operationalized using a small set of measurable goals for organizational input/output impacts such as the reduction of energy consumption, reduction or reuse of waste, and reduction of CO₂ contributions. As these goals are opposed to the pressure on organizations toward growth and an overall increased, albeit more efficient, use of resources (Meadows, 1998). Thus, the enterprises would need to implement the green view into their business vision making sustainability initiatives a potential issue for enterprise strategy management to stress its importance.

While this implication remains to be proven, the goal of this paper is to analyze the role that BPM plays in the realization of business sustainability initiatives, i.e. to explore the types of BPM techniques used in sustainability initiatives across industries. Defining this role will allow structuring, introducing and managing sustainability-focused business activities for C-level managers, business process owners and consultants. Therefore, 78 case studies of enterprises that have implemented sustainability or so called green activities into their business are analyzed. As the case studies were derived from the internet and their content was analyzed according to defined criteria, the research method applied here is the online content analysis. The contribution of this paper is twofold: its results add to the discussion on the realization of enterprise sustainability through BPM as well as illuminate the origins of BPM and its relation to enterprise strategy realization for enterprise-wide sustainability management. The paper is structured as follows: First related work on enterprise sustainability and sustainable business process management is briefly reviewed in section 2; the research method and its limitations are presented in section 3. Results of analysis are then described in section 4. Discussion of the findings and outlook on further research towards the definition of BPM role for sustainable business operations finish the paper.

2 Related Work

Realizing potentials for a more sustainable business has been a research topic for over a decade now. (Loeser 2013) noticed the particular meaning of the goal definition identified by (Hart 1995) to advance the environmental sustainability of businesses: 1) pollution prevention, achieved through minimization of waste and emissions; 2) product stewardship, addressed by consideration of stakeholder demands and optimization of product lifecycles; 3) sustainable development, accomplished through a reduction of the organization's environmental footprint and commitment to a long-term sustainability vision. This definition of the stages of the environmental strategy is also considered in this paper to assess the strategic maturity of an enterprise.

BPM related aspects that focus on sustainability, i.e. green BPM, as a research topic is increasingly found on the ISR agenda. Opitz, Krup, and Kolbe (2014) define green BPM as “the sum of all IS-supported management activities that help to monitor and reduce the environmental impact of business processes in their design, improvement, implementation, or operation stages, as well as lead to cultural change within the process life cycle.” Process metrics play an important role in the assessment of the application results of BPM tools. Reiter, Fettke, and Loos (2014b) introduce a combined approach of IT and BPM for efficient energy use in a process. Cleven, Winter, and Wortmann (2012) discuss the capabilities required to measure and manage sustainability performance on a process level by providing a capability maturity model (CMM) for green process performance management capabilities. Goldkuhl and Lind (2010) address the process design phase by presenting an extended process modelling approach for capturing and documenting the greenhouse gas (GHG) emissions produced during the execution of a business process as well as an accordant analysis method. The calculation methods for the carbon footprint of a process already being explored by e.g. (Recker, Rosemann, and Gohar, 2011; Grimm, Ereik, and Zarnekow, 2013; Pan and Kraines, 2001; Heijungs and Suh, 2006; Cooper and Fava, 2006). Among these works, two general measurement approaches can be distinguished: Cooper and Fava (2006) suggest a bottom-up approach from the process analysis perspective, while Pan and Kraines (2001) describe a top-down perspective incorporated in the environmental input-output analysis. Heijungs and Suh (2006) combine the two approaches. Nowak et al. (2011) present a methodology and architecture for green BPR, providing a starting point for green process analysis and re-design. Further methods for process analysis towards environmental potentials are presented in, e.g. Hoesch-Klohe and Ghose (2012).

3 Research Method

To answer the research question about the role of BPM within sustainability initiatives case studies of firms that have already initiated activities to incorporate sustainability into their business have been identified and analyzed. The case studies published online by the performing enterprises themselves or by the consultancies build the set of analysis. To identify the case studies, an internet search using an internet search engine with different sets of keywords was performed. First the keywords “sustainable business case study” were used. The resulting cases have been pre-analyzed with the focus on the information on the realization activities, methods and their adherence to BPM. Cases that did not contain this information have been omitted from the further analysis. The search was repeated using the keywords: “Green initiative case study”, “enterprise sustainability case study”, “sustainable enterprise case study” and “green business process management / BPM case study”. After the pre-selection according to the criteria mentioned above in 78 cases from the years (approx.) 2004- 2013¹ were included for the analysis. The final analysis was performed on these cases using online content analysis method as the cases were found and managed online. The validity of the (online) sources and thus the reliability of the information were proven by researching the publishing source and homepage provider separately. To avoid possible changes in the case study descriptions, the cases were downloaded from the source and stored offline for further analysis. The content analysis was conducted using categories derived during the pre-test. These categories included information on the driving forces behind the sustainability or green initiative, including motives for the case studies and initiating sources within the enterprise. The level of the (potential) changes induced by the initiative within the enterprise (process vs. resource) was in the focus of the analysis. Other aspects of analysis were management methods used to realize the change as well as the effects that the changes induced on the enterprise elements, such as organization, strategy, processes, etc. For a compact representation, the enterprises were grouped in industries. The results of the analysis are shown in tables 1 and 2.

The research method applied here has its limitations. First and foremost secondary sources, i.e. documents describing the transformation on different levels, were used to derive the results. By the rather broad definition of the analysis criteria this challenge was addressed and limited. Also, 78 enterprises with specific characteristics such as: mostly small and middle ones (SME), situated in Australia, Ireland and USA, predominantly from the manufacturing industry (37.2%), were analyzed. The analysis results were derived using content analysis of case study descriptions, thus limiting the amount of specific details. Therefore, this study is exploratory in its nature and its results need further evaluation and affirmation using the empirical tool set. Nevertheless, they provide useful results for BPM researchers and practitioners interested in the implementation of sustainability within a process or enterprise.

4 Results of the analysis

The enterprises in the sample set were grouped in industries, resulting in ten industry categories: utility provider (three entities); manufacturing (29 entities), food (eight entities); services (eight entities); entertainment (including shopping, six entities); non-profit organizations (including government and non-governmental institutions, four entities); print (four entities); IT (including software development, data centres, and IT service providers, 12 entities) as well as education and research (four entities). All of the sustainability case studies were initiated top-down, i.e. as a program or a project, from the C-level management of the enterprise. This fact is also reflected in the goals of the case studies (see table 1), whereas the methods for the achievement of the goals are described in table 2.

¹ The case studies are often dated using the publishing date of the text, omitting the actual timeline of the project.

| Goal | Goal/ Industry | Services | Non-profit | IT | Entertainment | Print | Food | Manufacturing | Government | Education | Conglomerate | Utility |
|------|---|----------|------------|----|---------------|-------|------|---------------|------------|-----------|--------------|---------|
| A | Env. Aspects: reducing waste, resource usage, GHG emissions | 7 | 4 | 2 | 2 | 4 | 3 | 15 | | | | |
| B | Improve operations | | | 2 | 3 | | 1 | 6 | 1 | 1 | | |
| C | A and B | | | 3 | 3 | | 1 | 3 | | | | 1 |
| D | Economic growth without incrementing A | | | 3 | | | | | | 2 | | |
| E | Cost savings | 1 | | 1 | 1 | | 2 | 4 | | | | |
| F | A and E | | | 1 | | | 1 | | | | | |

Table 1. Industry and goals

Providing more sustainable processes was only set as a goal of a sustainability initiative once by a utility provider.

Table 2 shows the top goals for the case studies, i.e. the goals that were set by the most enterprises in the sample set, and the methods that were used to achieve these goals. In both cases multiple entries were possible. Especially in the enterprises from the manufacturing domain business analysts saw the benefit of Lean Manufacturing (LM) methods not only in enhanced product quality and cost savings. The resulting optimization of resource use and process optimization have also positive effects on environmental metrics such as energy and water use savings, waste reduction and the consequent reduction of GHG emissions that were not anticipated.

The involvement of LM and Six Sigma methods into the sustainability initiatives includes workshops with process owners and workers, enabling an increased awareness of the importance of sustainability efforts and optimization of the resource usage. Also, some companies joined a governmental program to design the strategy and actions towards sustainability (5.1%), while others engaged sustainability consultancies to analyze the status quo and potentials for sustainable development in the business operations (9%).

| Method | Goal | A: Env. Aspects: reducing waste, resource usage, GHG emissions | B: Improve operations | C: A and B |
|--|------|--|-----------------------|------------|
| Measuring/ Tracing consumption | | 9 | | 4 |
| Reduce/rethink resource usage | | 12 | | |
| Process redesign | | 3 | | |
| Employee training/ communication | | 4 | | 2 |
| Audits | | 5 | | |
| Tracking consequences/savings, Reports | | 2 | | |
| Recycling | | 5 | 2 | 1 |
| CIP/LM | | 3 | 4 | |
| Consultancy | | 3 | 2 | |
| Process automation | | 3 | 2 | 1 |
| Performance analysis | | 2 | | |
| EMS | | 3 | | |
| SCM | | 4 | | |
| Process analysis | | 1 | | |
| Kaizen | | 2 | 3 | 2 |
| Resource optimization/efficient usage | | | 5 | |

| | | | |
|-------------------|--|---|---|
| Lean environment: | | 2 | |
| Process changes | | 2 | |
| VSM | | 3 | |
| Process measuring | | 1 | |
| Process mapping | | 1 | |
| Consultancy | | 2 | |
| Benchmarking | | | 2 |

Table 2. Top goals and methods for their achievement

Among the sample set, the most used method (multiple methods were used in general) to meet a sustainability goal was to re-think and re-design the general usage of resources such as electricity usage for lightning and cooling of the facility, water usage and waste reduction. 30.8% of the enterprises used this method to meet their sustainability goals.

17 enterprises, i.e. 21.8%, put an explicit focus on process management methods. As Kaizen methods, value stream mapping (VSM) and continuous improvement process (CIP) are borrowed by BPM (Conger 2010) they are not included in the count of process-specific methods for sustainability realization. Among the BPM methods that were used for managing the end-to-end processes, process automation is the most deployed method (used by seven enterprises, i.e. by 9%), followed by process changes, including process re-design. Only six enterprises, i.e. 7.7% of the sample set, were ISO 14001 and two, i.e. 2.6%, ISO 9001 certified at the time the case study was conducted. Accordingly, only few enterprises (four, i.e. 5.1%) use an environmental management system (EMS) in their business. The methods of: Key Performance Indicators, process redesign, Continuous Process Improvement; Process virtualization; Industrial ecology method were each chosen once to fulfil the goal C.

The stages towards a sustainable enterprise defined by Hart (1995) were presented in section 2. 70.5% of the enterprises that were analyzed in the sample set named reduction of waste or GHG emissions explicitly as their goal for the sustainability initiative, positioning them on the first level of the sustainability maturity. 11.5% of the enterprises considered product optimization and stakeholder, i.e. customer, public or supply chain (SC) partner, involvement into their sustainability program. The third stage of a sustainable enterprise as defined by Hart, i.e. developing a long-term strategy or vision of the sustainability incorporation into their operations, was aimed for by 5.1% of the enterprises within the sample set. Other companies named cost reduction, operation efficiency, green product development or benchmarking for business strategy development among the goals of their initiative. Moreover, the methods that were applied within the case studies were rarely designed for the identification of sustainability aspects. Six enterprises, i.e. 7.7% of the sample set, used a specific method for sustainability metric management, such as Lean and Green or industrial ecology.

5 Discussion of the findings

Findings from this case study analysis indicate that the business community does not seem to have agreed upon a general set of reasonable and practical approaches for an efficient implementation of sustainability issues, as has already been discussed by (Labuschagne, Brent, and van Erck, 2005). Also, the majority of the sustainability initiatives focus on reducing the general resource usage such as electricity stimulated by the expected cost savings. It is noticeable that the initiatives were introduced by the C-level management and promoted through enterprise-wide communication of the achievement, accompanied by employee trainings (in five cases, i.e. 6.4%), while public or market pressure played only a minor role in the decision to involve into sustainability. For the identification of (resource) efficiencies, Kaizen events and CIP initiatives were often used mostly by enterprises from the manufacturing domain that have already implemented Lean Manufacturing or TQM techniques in their processes. A possible conclusion can be that management and implementation techniques for e.g.

TQM and sustainability seem to rely on similar mechanisms. This assumption however constitutes an exciting research topic (Dües, Tan, and Lim, 2013).

Cost savings was the second exclusive goal mentioned by the enterprises from the sample, implying that the environmental benefits that result from the accordant activities are considered as a byproduct of lean or optimization actions rather than the goal itself, while providing a unique proposition to gain customers and market share. Measurement and tracking of the resource usage has shown to be a frequently established first step in the sustainability initiative and also claimed to provide first insights on the importance of the green changes, e.g. in (Powell, 2009). Thus, resource and business activity monitoring and analysis of resource usage can provide a valuable tool for the realization of the sustainability strategy and awareness. Process management techniques, specifically techniques for process optimization, are also shown to result in environmental benefits, i.e. resource usage or waste reduction, without being explicitly focused on designing green processes.

As various industries are present in the sample, indications about favoured managing techniques for green initiatives among the industries can be deviated. It is noticeable that manufacturing companies tend to adopt lean and sustainable benefits but also that service-oriented enterprises financially and environmentally benefit from conscious resource usage by applying and adopting same techniques. Thus, using BPM in such projects is probable to result in partial and limited impacts instead of promoting the long term sustainability vision within the enterprise.

6 Summary and Outlook

In this paper, an analysis of real-life sustainability initiatives was conducted. Results of this explorative study offer the conclusion that BPM tools and approaches can provide a basic support for the enterprise transition towards sustainability if they are initiated and supported by upper management. The findings of the case study analysis as well as the research findings investigating strategy implementation implicate that a bottom-up, i.e. process-based, realization of sustainability goals will probably not result in a more sustainable organization as implied in (Pernici et al., 2012), as basic change management techniques need to accompany a successful realization (Trkman, 2010) and integration of the green thinking into the business strategy. This view is also supported by the ISO 14001, where the implication is that the environmental policy is set in the beginning of the Plan-Do-Check-Act cycle.

Future work on the definition of the role of BPM in the transformation of a firm towards sustainability will include the exploration and evaluation of the already suggested potentials of BPM life cycle and methods with sustainability factors as well as conducting case studies and expert interviews to show the plausibility of the top-down approach for a business transformation towards sustainability.

References

- Brundtland Commission. 1987. *Report of the Brundtland Commission of the United Nations: Our Common Future*. Oxford, UK.
- Cleven, A., Winter, R., and Wortmann, F. 2012. "Managing Process Performance to Enable Corporate Sustainability: A Capability Maturity Model." In *Green Business Process Management Towards the Sustainable Enterprise*, 111–29.
- Conger, S. 2010. "Six Sigma and Business Process Management." In *Handbook on Business Process Management 1*, 127–48. Springer Berlin Heidelberg.
- Cooper, J. S., and J. A. Fava. 2006. "Life-Cycle Assessment Practitioner Survey: Summary of Results." *Journal of Industrial Ecology* 10 (4): 12–14.
- Dües, Ch., Tan, K., Lim, M. 2013. "Green as the New Lean: How to Use Lean Practices as a Catalyst to Greening Your Supply Chain." *Journal of Cleaner Production* 40 (February): 93–100.
- Goldkuhl, G., and Lind, M. 2010. "A Multi-Grounded Design Research Process." In *DESRIST'10*

- Grimm, D., Ereğ, K. and Zarnekow, R. 2013. "Carbon Footprint of IT-Services – A Comparative Study of Energy Consumption for Offline and Online Storage Usage." *AMCIS 2013*
- Hart, S. L. 1995. "A Natural-Resource-Based View of the Firm." *Academy of Management Review* 20 (4): 986–1014.
- Heijungs, R., and S. Suh. 2006. "Reformulation of Matrix-Based LCI: From Product Balance to Process Balance." *Journal of Cleaner Production* 14 (1): 47–51.
- Hoesch-Klohe, K. and Ghose, A. 2012. "Environmentally Aware Business Process Improvement in the Enterprise Context." In *Harnessing Green IT: Principles and Practices*, 265–82. Wiley.
- Labuschagne, C., A. C. Brent, and R. P. G. van Erck. 2005. "Assessing the Sustainability Performances of Industries." *Journal of Cleaner Production* 13 (4): 373–85.
- Loeser, F. 2013. *Green IT and Green IS: Definition of Constructs and Overview of Current Practices. AMCIS 2013*.
- Loos, P., Nebel, W., Gómez, J., Hasan, H., Watson, R., Brocke, J., Seidel, S. and Recker, J. 2011. "Green IT: A Matter of Business and Information Systems Engineering?" *Business & Information Systems Engineering* 3 (4): 245–52.
- Meadows, D. 1998. *Indicators and Information Systems for Sustainable Development*. Hartland: Sustainability Institute.
- Molla, A., V. Cooper, and S. Pittayachawan. 2009. "IT and Eco-Sustainability: Developing and Validating a Green IT Readiness Model." In *ICIS 2009*, 141. <http://works.bepress.com/siddhi/4/>.
- Nowak, A., Leymann, F., Schumm, D. and Wetzstein, B. 2011. "An Architecture and Methodology for a Four-Phased Approach to Green Business Process Reengineering."
- Opitz, N., K. Ereğ, T. Langkau, L. Kolbe, and R. Zarnekow. 2012. "Kick-Starting Green Business Process Management–Suitable Modeling Languages and Key Processes for Green Performance Measurement." In *AMCIS 2012*.
- Opitz, N., H. Krup, and L. M. Kolbe. 2014. "Green Business Process Management -- A Definition and Research Framework." In *2014 47th Hawaii International Conference on System Sciences*, 3808–17. IEEE. doi:10.1109/HICSS.2014.473.
- Pan, X., and S. Kraines. 2001. "Environmental Input–output Models for Life-Cycle Analysis." *Environmental and Resource Economics* 20 (1): 61–72.
- Pernici, B., Aiello, M. and Donnellan, B. 2012. "What IS Can Do for Environmental Sustainability : A Report from CAiSE ' 11 Panel on Green and Sustainable IS.", 30 (18).
- Powell, J. 2009. "Fair Isaac." *Enterprise Systems Journal* 3(3).
- Recker, J., M. Rosemann, and E. R. Gohar. 2011. "Measuring the Carbon Footprint of Business Processes." *Business Process Management Workshops* January: 511–20.
- Reiter, M., Fettke, P. and Loos, P. 2014. "Towards Green Business Process Management: Concept and Implementation of an Artifact to Reduce the Energy Consumption of Business Processes." *HICSS 2014*, January. Ieee, 885–94.
- Trkman, Peter. 2010. "The Critical Success Factors of Business Process Management." *International Journal of Information Management* 30 (2): 125–34.
- Vom Brocke, Jan, Stefan Seidel, and Jan Recker, eds. 2012. *Green Business Process Management*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-27488-6.
- Wesumperuma, A.; Ginige, A.; and A. Hol. 2011. "A Framework for Multi-Dimensional Business Process Optimization for GHG Emission Mitigation." In *AMCIS 2011*.

DESIGN OF A NEW DATA STRUCTURE TO SUPPORT NON-INVASIVE DIAGNOSTIC ON HERITAGE METALS

Completed Research

Rosselet, Antoine, University of Applied Sciences Western Switzerland (HES-SO), HEG Arc, Neuchâtel, Switzerland, antoine.rosselet@he-arc.ch

Rochat, Vincent, University of Applied Sciences Western Switzerland (HES-SO), HEG Arc, Neuchâtel, Switzerland, vincent.rochat@he-arc.ch

Gaspoz, Cédric, University of Applied Sciences Western Switzerland (HES-SO), HEG Arc, Neuchâtel, Switzerland, cedric.gaspoz@he-arc.ch

Abstract

Conservation of heritage artifacts is a very sensitive task as conservators usually have very little information about the artifacts. Moreover, due to the uniqueness and the historic value of these artifacts, invasive analysis are not always possible. Therefore, without sampling options, conservators are required to use non-invasive diagnostic methods in order to identify the metal characteristics of the artifact. When confronted with an unknown artifact, conservators generate conceptual models of the corrosion forms. These models are based on formal representations of corrosion forms, but are not directly exploitable for drawing hypotheses regarding the underlying metal. This paper presents the design of a data structure generated from the conceptual models which supports the comparison and retrieval of corresponding artifacts. Integrated with a database of heritage artifacts, this data structure offers advanced decision support to conservators confronted with unknown artifacts.

Keywords: conceptual model, data structure, design science, decision support system

Introduction

The restoration of heritage artifacts requires careful identification of the material before any treatment can be performed. Over the years, the field of conservation-restoration of ancient or patrimonial objects evolved toward better preservation of the artifacts, mainly by reducing invasive analysis and applying very specific treatments in order to avoid further deterioration. These changes imply two assumptions. First, in order to reduce or totally abandon invasive analysis of the material, such as sampling, we need to be able to perform the same diagnostic with non-invasive techniques such as microscopic analysis. Second, in order to apply specific and targeted conservation treatments, restorers need to identify univocally the composition of the artifact.

In this paper, we present an innovative way of transforming non-invasive analysis of ancient metallic artifacts in actionable data structures that can be processed, compared and analyzed in order to predict the metallic composition of the artifacts. Currently, the most used non-invasive technique of heritage metals identification is stratigraphic analysis (Bertholon, 2000). This technique borrows from the field of geology and archeology by representing corrosion forms as a superposition of layers also called stratigraphies. After many years buried in soil or water various corrosion products and deposits (for example sediments) affect the metal contained in the artifact. The restorer visually observes these corrosion products and deposits in order to construct a representation of the different layers of deposits and corrosion products, like a geologist would do with a core sample.

Currently, there is no support tool to assist conservators in building their stratigraphies, mainly due to the novelty of the method. It is mainly a pencil/paper method, relying on defined construction rules (Figure 1). In order to support the broader diffusion of the method, its initiators are looking for a tool to assist conservators in using it. However, even if the term is not used by the authors of the method, the analysis of corrosion products using stratigraphies is some sort of conceptual model of the analysis of historic corroded metals. However, we can then build on the rich experience of conceptual modeling and data structures to design a tool to assist restorers in identifying underlying materials in conservation projects on historic artifacts.

In this paper, we start by introducing the notion of conceptual modeling and show how it is related to the stratigraphic methods developed by Bertholon (2000). Then, we draw on the literature on data structure to show how we use the generated conceptual model to produce a digital model (data structure) of the corrosion products of the artifact. We then explain how we designed the digital model and how we translated the conceptual model into a graph model that can be used to analyze and compare stratigraphic representations from various artifacts. Finally, we present some scenarios in which the data structure can also be used to increase its value for professionals in the conservation-restoration field.

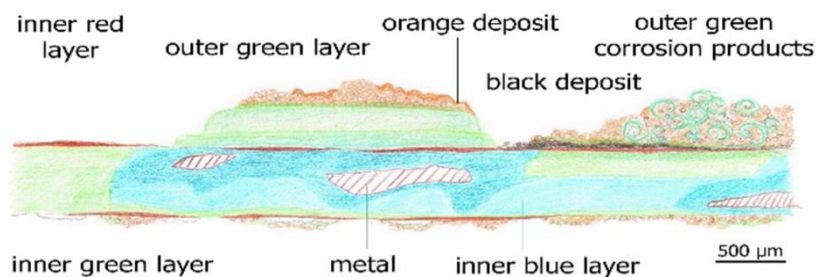


Figure 1. A stratigraphy, as drawn by a conservator

The principal contribution of this paper is a new data structure that can be used to store stratigraphies of heritage artifacts in a graph database in order to support professionals in identifying the metallic composition of these artifacts without invasive sampling. The subsidiary contribution is a design pro-

cess that can be used to tackle similar problems that need to generate digital representations of analogue models in order to process them in some way.

From reality to data structure

Conceptual models are widely used in the information systems (IS) field (Wand & Weber, 2002). Kung and Solvberg identified their main purposes: (1) supporting communication between developers and users, (2) helping analysts understand a domain, (3) providing input for the design process and (4) documenting the original requirements for future reference (as cited in Wand & Weber, 2002). In management science, models are defined as “a representation of reality intended for some definite purpose” (Pidd, 2010). These definitions contain two elements that are related to our research objective. First, because we are not able to sample the observed artifacts, we will not be able, during the creation of the conceptual model, to univocally identify any parts of the reality. So, we will have to rely on observations of the artifacts in order to create a stratigraphy representing its corrosion products. Second, the model is created to achieve a specific goal. We are not trying to have a model representing the shape, color, texture or hardness of the object. Rather, we are seeking to create a model that will contain sufficient information to help us determine the nature of the metal composing it. In order to support these goals, the conceptual model should abstract an appropriate simplification of the reality.

Furthermore, the conceptual model needs to be designed so that it can later be translated into data structures that are relevant to our goal and which will help us understand the problems we want to address. Indeed, many different conceptual models can be made from a same reality. Depending on our needs, some conceptual models are pertinent whereas others are not. This is why an analysis phase needs to be undertaken in order to design an appropriate and actionable conceptual model that can further be used (Wand, Monarchi, Parsons, & Woo, 1995).

One well-known model is the subway map that a public transport operator would produce. The map is a simplified representation of the physical network, which allows passengers to work out possible routings through the network of lines and stations. However, these maps deliberately distort the reality in order to support their purpose. Lines are distorted in order to emphasize their general direction and the interchanges, regardless of their physical layout. This partial representation of the reality fulfills its goal of supporting passenger routing, but is of no use for contractors dealing with network maintenance. Pidd has a more specific definition of the model that we will use in the rest of this paper. “A model is an external and explicit representation of part of reality as seen by the people who wish to use that model to understand, to change, to manage and to control that part of reality.” (Pidd, 2010)

As we saw in the introduction, the analysis of corrosion forms on ancient artifacts requires specific knowledge of its constituent corrosion products. In order to be able to compare, analyze and discuss these corrosion products, we need to create a stratigraphic representation of them. According to Robinson (2010), the first model to create and which is based on the reality we want to understand is called the “conceptual model”. It is the first step which can then lead to a data structure stored in a computer. Once stored, the data structure can be used to perform simulations and to find out behavior patterns from the real world. In our context, we want to use conceptual models (1) as a final product to have a visual representation of an artifact to work on and to talk about and (2) as an intermediate step linking reality to data structure.

To illustrate the first use, we can think of the business model canvas (Osterwalder, Pigneur, & Tucci, 2005). We start with the business we want to analyze (reality). From the structure and functioning of this business we can draw a canvas, which is a representation of the reality. It consists of a conceptual model and a common tool to convey and discuss ideas about the business. The model can also lead to the design of new business models. Another example is the design of a Business Process Diagram (BPD). Its elaboration starts with an observation of the real world—in this case how people are orga-

nized within an organization—and results in a diagram that follows the rules dictated by the Business Process Model Notation (BPMN) (White, 2004). People can then discuss, analyze and develop that model. As it stands, the model is sufficient to support its purpose without needing to translate it into a computer language.

Note that in the two examples above, the conceptual model generated followed the rules dictated by the canvas or the framework in which it operates. However, this is not necessarily the case. The main advantage of modeling within a framework is that people only need to know how the framework works to understand the model and therefore the real world it depicts. It also allows easy comparison of different models, as they use the same notation conventions.

The second use is what really interests us. Indeed, as we want to automate the process of stratigraphic representation, we need a computer system to help the user construct the stratigraphy and to store it in a database. A simple drawing cannot be directly used to run simulations or to perform further tasks. Therefore, we need a data structure representing the stratigraphy and supporting further processing (Peuquet & Marble, 1990).

From reality to conceptual model

In order to build an actionable data structure, we first need to understand how to translate the perceived reality in a conceptual model that can be shared between, and understood by, users. When designing electronic circuits, engineers start from a conceptual model showing the different components of the electronic circuit. Each component and its characteristics are represented by specific drawings in the model. This helps the engineer understand the circuit structure and to compare it with other circuits. This can also be used to compute variations in current along the circuit. The same analogy can be made in the context of workflow management. We start by modelling the main process supporting the workflow using common notations such as BPMN. Here again, the BPD is a conceptual model of a real process taking place in a company. One particularity of this model is that it does not look like reality. Indeed, we do not “recognize” a human activity when a rounded rectangle is drawn; it only looks like a rectangle, not like a human performing a task. The same is true for most of the BPMN components. However, someone who knows the notation conventions can easily understand the model and the real process beneath it. An effort of abstraction is needed, but once it is done, the process is understandable. Manual comparison between different BPDs is also made possible thanks to BPMN.

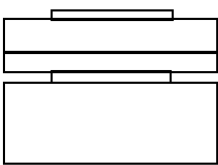
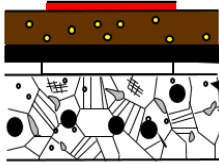
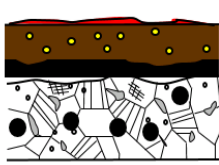
| Structure (one thick crust + metallic core) | Corrosion form (with characters of strata) | Corrosion form (with characters of strata / interfaces) |
|--|---|--|
|  <p>CP1e CP2i CP3i SV M</p> |  <p>CP1e (Fe,O) CP2i (CuFeS₂) CP3i (Sn,O,Cu) SV Pb Cu₂S M (Cu₈Sn_{1.5}Pb)</p> |  <p>CP1e CP2i CP3i SV M</p> |
| <p>Coding: CP1e;CP2i;CP3i;SV;M Level of accessibility: basic</p> | <p>Corrected coding: CP1e;CP2i;CP3i;SV;M Level of accessibility: advanced</p> | <p>Correct location of strata Level of accessibility: final</p> |

Figure 2. Conceptual models of stratigraphies

In our context, conservators will use specific notations to represent their visual observations. There is no wish to create a truthful representation of the artifact (like a map), but to create an abstraction of the reality in order to learn and share findings about the specific assembly of layers of corrosion products. Therefore, the conceptual model focuses on the area of interest for conservators: layers of corrosion products and their interfaces, as presented in Figure 2.

From conceptual model to data structure

Conceptual models are useful for sharing, studying and analyzing a specific view of the underlying reality. However, they are of limited use when computational processing is required in order to gather more information than the simple abstraction of the reality. Thus, if we want to apply computational processing to the conceptual model for further analysis, we need to transform it to a format that can be understandable to a computer. The translation is not always straightforward; refinements may be needed, depending on the conceptual model we use.

We have seen before that conceptual models can be used to represent the flow of current in a circuit. However, these models can only be used as a basis for understanding the functioning of the circuit. They are of little use in understanding what happens if a specific part of the circuit is altered. In order to comprehend the dynamics of the circuit, various software packages have been developed (Cellier & Clauss, Christoph, 2007) to support users in designing a circuit with building blocks—the blocks used in the conceptual model—and to derive equations from the model. These equations can then be used to simulate the model behavior, which helps deduce how the real electronic circuit works. The same applies to BPMN. In order to test the flow of the process we need a new language, the Business Process Execution Language (BPEL) (White, 2005), that can be executed in order to follow, step by step, the execution of the process. Therefore, BPEL gives us the ability to translate the conceptual model (a set of shapes) into an executable workflow.

However, many steps are needed in order to perform such a translation. When translating BPMN to BPEL, the BPD—made up of BPMN elements—first needs to be broken down into smaller parts. In this case, the graph structure is transformed into a block structure, with one entry point and one exit point for each BPD component. These parts are then individually translated into BPEL, depending on their type. For instance, a component that is identified as a sequence will be translated into BPEL with a sequence tag, whereas a component listed as a switch will display a switch tag. Each tag has its own characteristics, sub-characteristics and behavior. Therefore, business rules are taken into account, as a switch component does not behave like a while component and is composed of different elements. It is easier to respect these rules when the process is broken down into many components, as each one is individually analyzed and translated (Ouyang, Dumas, Aalst, Hofstede, & Mendling, 2009).

Corrosion product stratigraphies are made of superposed blocks representing a specific corrosion product layer. These layers can have characteristics such as color, hardness, friability, inclusions etc. Up to a point, these characteristics can be embedded in the representation of the strata. However, in order to retrieve similar stratigraphies, we need to be able to transform these graphical representations of corrosion products into an executable representation. A simple image analysis is not sufficient as we are interested in comparing the individual characteristics of the strata and not their appearance.

Contextually useful data structures

Finally, due to the fact that conceptual models are by definition simplifications of the reality, we need to find a way to add detailed information on parts of the models where information cannot be explicitly represented by the modelling conventions. For instance, in BPMN there is no notion of timing, or for how far the employees performing the tasks are from each other. In a map, we will generally not indicate the weather. Of course, some of these elements are not important. In fact, it depends on how we want to use the model; it depends on the data structure we need. As well as providing a common framework to work on and to talk about, the conceptual model allows us to imagine the visualization of the data structure; it becomes a prototype for creating computer solutions. When developers are asked to design software representing the data structure of a reality, they need to know the needs of the people who will use it and how they want to see their data—that is why it is so important in the process which leads to data structures.

In order to generate the appropriate data structures, it is essential to add some elements that we lost during the translation of the reality into a conceptual model. Indeed, as it is said for the BPMN: “the diagram itself will not display all the information required to create a valid BPEL file. A diagram with all that information would be too cluttered to be readable. A BPMN diagram is intended to display the basic structure and flow of activities and data within a business process. Therefore, a modeling tool is necessary to capture the additional information about the process that is necessary to create an executable BPEL file.” (White, 2005)

The data structure needs to include this additional information in order to respect the business rules. White (2005) adds this information when generating a BPEL using a dedicated tool which knows what kind of extra elements are needed when translating BPMN to BPEL. Thus, the modeler has to give the right information to the tool so that the latter can create a correct BPEL file that can later be used for various applications. When the design of a BPEL is finished, it is then possible to analyze its structure and validate the underlying process after some modifications are applied. Schmidt and Stahl suggest transforming BPEL processes into Petri nets to do so. Therefore, it would be possible to verify properties, such as checking whether a customer will always get an answer in the analyzed process (Schmidt & Stahl, 2004). In order to generate data structures, we started from reality by going through conceptual data. Eventually, when we have an exploitable tool that can validate an existing process, the path has been reversed: based on the data structure, we can assess a real process. Thus, this assessment can be very useful from the perspective of a process enhancement or a deeper analysis.

On the one hand, it is essential to create a conceptual model from reality, so that we can come up with data structures based on this model. On the other hand, depending on the information we are looking for, it is important to add more elements to the data structure than the ones that are embedded in the conceptual model. This will provide us with data that can then be manipulated in order to perform further tasks, such as simulation. This process is called “supercharging”.

Non-invasive diagnostic on heritage metals

The conservation of ancient and historic metal artifacts requires a detailed understanding of both their composition and their alteration. Therefore, conservators have developed preventive conservation strategies and curative treatments in order to stabilize active corrosion processes, while preserving all relevant information. However, the invasive and/or destructive character of metallography and some associated chemical analysis is the main factor which limits their application to cultural heritage artifacts. In the past, the fields of archaeometallurgy, history of techniques and conservation science have produced large quantities of metallographic studies using samples and/or physical and chemical analyses (invasive, non-invasive, destructive and semi-destructive). These studies cover all pre-industrial metal families (Fe, Cu, Ag, Au, Pb and Sn). Unfortunately, this data rarely contains information on corrosion forms and the nature of the corrosion products (Degrigny & Senn, 2012).

For this project, the stratigraphic representations were modelled using the methodology developed by Bertholon (2000). It is based on the fact that the structure of artifacts can be broken down into several layers (strata), which together constitute a stratigraphy. A stratum can be the metal itself or any corrosion products that have affected the metal. Each of these layers has several characteristics depending on its nature. The goal of the conservator is to find out the metal of which the artifact is made. In order to do so, layers have to be removed and treatment has to be applied. Being able to compare the artifact with other artifacts that have similar corrosion products and which are listed and analyzed in a report is highly valuable for a conservator, as this could give clues about how to treat it.

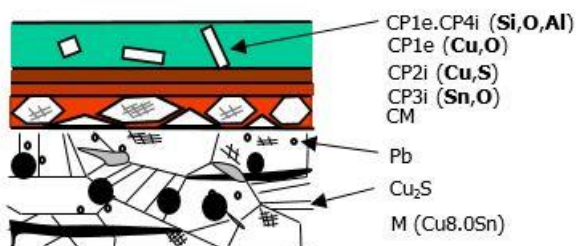


Figure 3. Example of a stratigraphy drawn on a computer

After different strata have been identified and manually drawn, they are reproduced in a computer, using special notation conventions as seen in Figure 3 (Bertholon, 2000). This constitutes the conceptual model. Then, the comparison is straightforward as all the artifacts drawn follow the same modeling framework. Nevertheless, the comparison is only visual; thus far, the stratigraphies have not been coded into a computer language that could compare them, based on objective characteristics.

Without a tool that can regroup the stratigraphies, the conservators have difficulty identifying the metal the artifact is composed of because comparing it with others is complicated. Therefore, applying any curative treatment is arduous. A remaining option would be to take an artifact sample. But as this is forbidden for most artifacts by professional ethics, the conservator has to find another way to identify the metal.

In order to apply the right restoration method to an ancient metal artifact without damaging it, the more information and comparison the conservator has at hand, the more effective the work will be. Therefore, providing the conservators with a tool they can use to construct and compare stratigraphies would be extremely helpful as it would facilitate their work in identifying the metal because a comparison would then be possible. There are also many advantages of having a tool that could be used by conservators from all around the world. New corrosion structures could be identified and added to the database. In time, a community of conservators could be created based on this methodology and this tool, which could then evolve over time. The database could then expand with more artifacts and stratigraphies.

Design of a data structure for stratigraphies

In order to support the non-invasive diagnostic of ancient metals we needed to design a suitable data structure that could be used to compare and analyze corrosion products, to support the conservators in identifying the composition of the metal. We started by analyzing the conceptual models used to represent the layers of corrosion products in order to propose a methodology for designing the data structure. We encountered several issues at this first stage: (1) there is no single source of information regarding the various characteristics of the stratigraphies, (2) there are hundreds of characteristics that need to be grouped in order to maintain the overview, (3) there are multiple representations of the same reality due to the fact that the methodology is relatively new and (4) there is currently no exhaustive information on the characteristics of these stratigraphies due to the lack of data.

The first step that we took was to start with a small ontology of the domain, in order to identify the main informational components of an artifact. We identified the 12 concepts listed in Table 1 and displayed as a graph in Figure 4.

| Concept | Usage |
|----------------|---|
| Artifact | Represents a heritage metal object |
| Environment | The environment to which the artifact was exposed (fire, water, soil, etc.) |
| Corrosion type | The corrosion type (Type II, etc.) |
| Technology | The technology used to craft the artifact |

| | |
|----------------|---|
| Owner | The current owner of the artifact (museum, etc.) |
| Object | The main use of the artifact at the time it was crafted (bracelet, etc.) |
| Origin | The geographic origin of the artifact |
| Chronology | The estimated period in which the artifact was crafted |
| Stratigraphy | The representation of all corrosion products of an artifact that could be observed at the same location on the artifact (Bertholon, 2000) |
| Stratum | One corrosion layer of the stratigraphy (Bertholon, 2000) |
| Characteristic | A property of a stratum (Bertholon, 2000) |
| Interface | The junction of two strata from the same stratum (Bertholon, 2000) |

Table 1. Main concepts of the ontology representing an artifact and its corrosion layers

Starting with the ontology, we identified the main characteristics of the strata from the available literature. We grouped these characteristics into families of characteristics, each with a description, a selection attribute indicating if the family can be used more than once in each stratum and a list of dependencies (for example the Corrosion Product Stratum Composition can only be used once on a single stratum and the silver metal can be related to chlorine or sulfur). At the end of this cycle, we identified 523 characteristics and their relations. Each characteristic of the model is a member of a family and has relations to other characteristics.

We then considered many different options regarding the methods and the database schemas to choose in order to store the stratigraphies. There are currently two types of database—relational and graph—that can be used to store such a data structure. The results of recent research tend to observe better performance with graphs stored in graph databases rather than in relational databases, even if both ways are well supported by current database systems (Batra & Tyagi, 2012). Graph databases are also more flexible in the case where new relationships need to be added, which is an important property of our own data structure. This led us to choose the Neo4j database (Neo4j, 2012) as the backend for our data structure. When storing data, every node in the database is either an artifact, a stratigraphy, a stratum, an interface (between two strata), a characteristic or a sub-characteristic. These elements are then connected via relationships. The usual pattern is to start from an artifact, which has one stratigraphy made up of several strata. Each of these strata has different characteristics and sub-characteristics, as shown in Figure 5.

The next step was to generate the data structures out of the conceptual model. As we have seen, in order to add information that is not represented in the conceptual model, we have to “supercharge” the model. This allows us to add the necessary elements to generate pertinent data structures. This step was done in collaboration with experts in conservation-restoration. We built a support system in order to aid the selection and application of specific characteristics to strata that were drawn from the stratigraphies. For each stratum, the expert had a choice of families to choose from, based on their relations with the type of stratum. Each time a characteristic is selected and based on its relations with other characteristics, the choice of families is adapted to reflect the current construct. At the end of this step, we had a graph database containing all the stratigraphies of the artifacts that we used as our test sample.

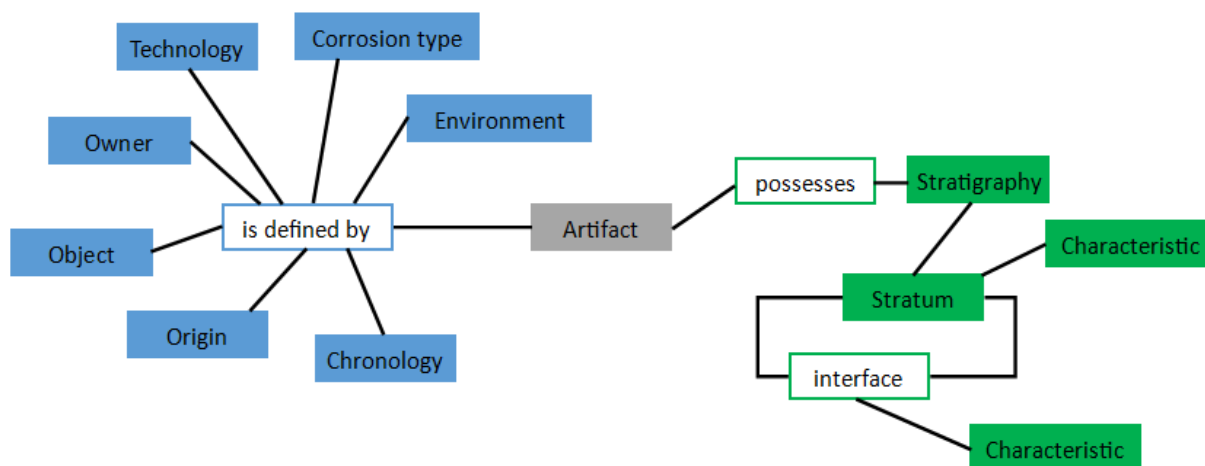


Figure 4. A graph representation of our ontology

We are now able to take a conceptual model of corrosion products (a stratigraphy) and store it in a graph database, supercharged with various families of characteristics, depending on the type of the stratum. The next step was to recall the stratigraphy from the database in order to assess its completeness. There are multiple query languages available to store and retrieve data from a graph database—in our case, the Neo4j database. The preferred approach is to use Cypher (“Cypher Query Language,” 2015), a query language specifically developed for Neo4j and offering the developer the same experience as SQL for relational databases (Holzschuher & Peinl, 2013). However, the overall performance of the queries performed with Cypher is currently not as good as using native access to the data provided by the database. In our case, there are no contraindications for using Cypher to query our data structure because the performance reduction compared to native access is negligible for our usage.

Based on the artifacts’ characteristics, we created a comparison index, which helped us determine the closest artifacts in terms of structure and relevance for the researchers in conservation-restoration. That index was refined many times after suggested adjustments made by the researchers. With this index, our tool is now able to return the closest stratigraphies to a given stratigraphy, supporting researchers in finding potentially similar artifacts to the one they are studying. However, currently all queries have to be made using Cypher queries, which is not really suitable for our intended audience of conservation professionals.

The final product of the data structure is a digital model presenting all families and their characteristics. It can be used by professional conservators to properly supercharge a stratigraphy with detailed characteristics and by a developer in order to build a rules-based information system that can support the creation of new stratigraphies.

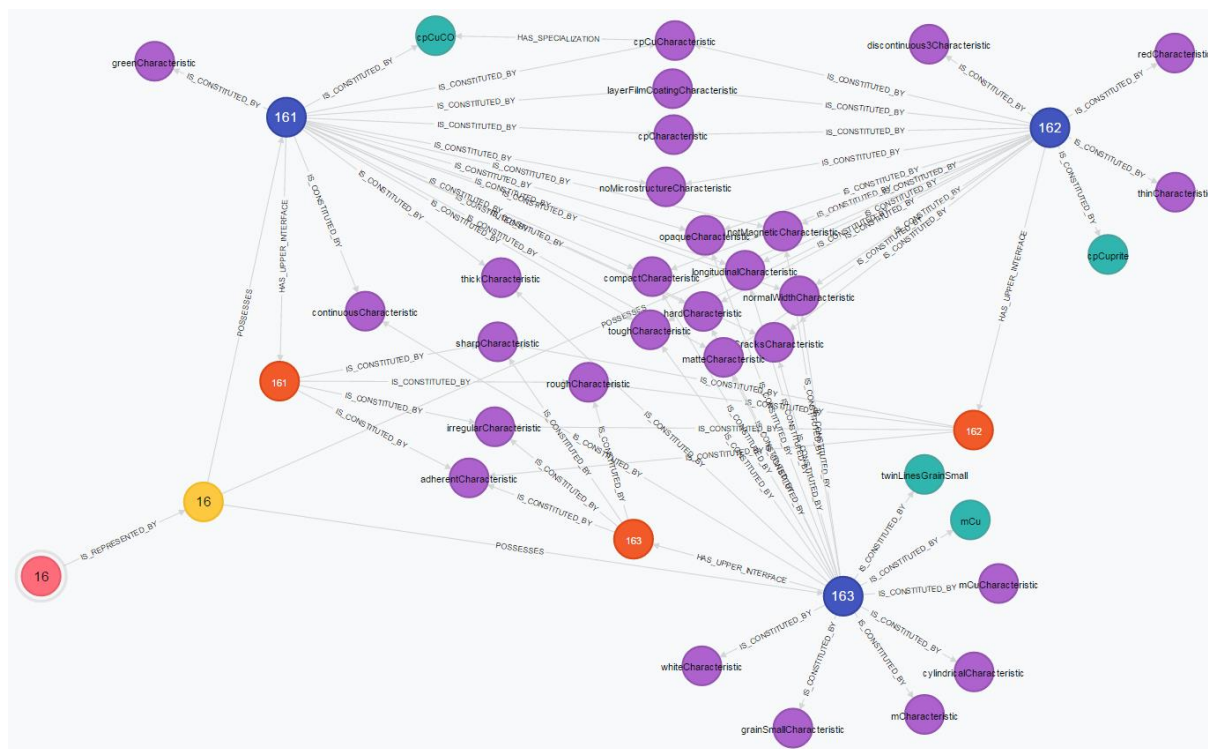


Figure 5. A graph representation of artifact 16 in Neo4j

Evaluation of the data structure

In order to evaluate the created data structure, we performed two types of tests. The first set of tests were aimed at checking the validity of the results returned by a comparison query and the second set were performed to assess the completeness of the data structures generated from the conceptual models and backward.

The comparisons of artifacts were performed using Cypher queries on the artifacts that were stored in our database. As seen before, the stratigraphies were created and checked by professionals of the domain. Artifacts data were also completed by information on their exact composition, so that we could assess the results of the comparison queries. We started to run comparison queries on the artifact itself. Hopefully, this returned a 100% matching result. Then we started to search for similar artifacts. Again, performing the query between one artifact and all artifacts stored in the database gave us a 100% matching between the artifact and itself, which was expected.

On the other hand, we found some non-correlated artifacts with exceptionally high matching. After carefully studying the results, we discovered that some typical corrosion products were neglected in favor of highly similar one and resulted in biased results. Common errors were found between artifacts in different environments. Artifacts buried in soil or in water have different corrosion products but can be made of the same metal. However, due to the relatively small sample of artifacts in the current database, we are not able to adapt our comparison queries at the moment. We did not have enough artifacts to identify all discriminant characteristics of our model. Nevertheless, returning more accurate results could easily be done by increasing the weighting we implemented for every characteristic.

In the following table, we compared artifact 16 with the others. As mentioned before, the result of this query returned artifact 16 as the best match. Next, artifact 17 is the second best match. Conservators confirmed that it is indeed the closest artifact to artifact 16 in terms of its composition and structure, which validates our comparison request.

| ID | # of strata | Difference # of strata | Total matching | Total relations | Matching100 |
|------------|-------------|------------------------|----------------|-----------------|-------------|
| artifact16 | 3 | 0 | 64 | 64 | 100% |
| artifact17 | 3 | 0 | 59 | 64 | 92% |
| artifact5 | 4 | 1 | 75 | 86 | 87% |
| artifact15 | 5 | 2 | 92 | 108 | 85% |
| artifact6 | 5 | 2 | 79 | 97 | 81% |

Table 2. Results of the comparison request with artifact 16

As a result, we are now able to return and sort the artifacts based on the matching with the artifact we want to compare. However, further work has to be undertaken in order to give more accurate results. In the future, we will have more and more stratigraphies in our database, which will help the conservators assess the pertinence of the results. This will also generate more cases in which we will need the conservators' expertise. Therefore, we will have to collaborate closely with them as the database expands.

The second part of the evaluation focused on the information that could get lost in the process from reality to data structure. The first loss is straightforward and always happens when designing a model: as the conceptual model is a simplification of reality, it does not include every aspect of the artifact which the model is based on. Afterwards, in order to have data structures as close to reality as possible, we had to supercharge the conceptual model with characteristics that were not easily understandable or deducible from the drawing. Furthermore, we were able to add annotations to the relationships between an artifact and its characteristics. However, some of these annotations are not taken into account when performing a comparison request. Indeed, a comparison request is only based on the characteristics to return the results and some annotations cannot be translated into characteristics, either because they only apply to a particular artifact or because they are too rare to be implemented in the data structure.

This leads to a second loss of information that occurs in the data structure. When adding an artifact stratigraphy into the database, the conservator has to choose between a limited number of characteristics. Indeed, researchers must provide information about the characteristic family and sub-characteristics for every new addition, whereas they can draw anything on a stratigraphy made by hand. This forces researchers in conservation-restoration to choose between a small range of characteristics and asks them to be very specific and to know exactly what their artifacts are made of. By contrast, they can draw new characteristics on paper, without being absolutely sure about what they are. The main advantage is that this formalism prevents researchers giving different names to the same characteristic. The comparison is therefore easier as the same pattern is followed. However, on our side we have to be as exhaustive as possible concerning the pool of characteristics that the researchers can select from when building their stratigraphies.

Table 3 sums up the completeness of our data structure compared to the conceptual model from the conservators. We can see that we still need to add some characteristics into our data structure to make them available for researchers when building their stratigraphies. Indeed, there are 92 more characteristics that should be in the data structure because they were identified by conservators as such.

| | | Data structure CR | | Total |
|---------------------|--------------|-------------------|--------------------|-------|
| | | Represented | Missing | |
| Conceptual model CR | Existing | 514 (correct) | 92 (to add) | 606 |
| | Non-existing | 9 (error) | undefined (future) | 9 |
| Total | | 523 | 92 | |

Table 3. Comparison between the conceptual model and the data structure

But there are 9 characteristics that are represented in the data structure but that were not pertinent for the conservators, either because we misinterpreted them or because we made a mistake when adding them to our database. As the database expands, we will certainly find that we have not identified all cases that will occur in the future. Therefore, the main focus will be on the attention paid to the correct identification of the characteristics in order for the data structure to be as accurate and as close to reality as possible. The closer the data structure is to reality, the more accurate the comparison requests will be, as the characteristics will be correctly identified.

Further research

At the moment, as we have to perform Cypher queries to be able to compare stratigraphies, only someone who knows how the Cypher language works and who can interpret the results returned by the comparison requests can make use of our tool. Therefore, we have to develop an intuitive tool that the researchers in conservation-restoration can use, as they are the end users. Of course, the tool must be based on our Neo4j database. This guarantees accurate results when comparing artifacts with one other.

If we want our tool to be used efficiently, we need it to be as close as possible to what the researchers and the students in conservation-restoration are accustomed to. Thus, as they provided us with their conceptual model, we immediately had the idea of using it as the visualization of our data structure. This adds a whole new set of possibilities. Indeed, the user would be able to create a stratigraphy that will be stored into our graph database. Based on its characteristics, the comparison with other stratigraphies will be easy. The user will also be able to visualize the stratigraphy while it is under construction.

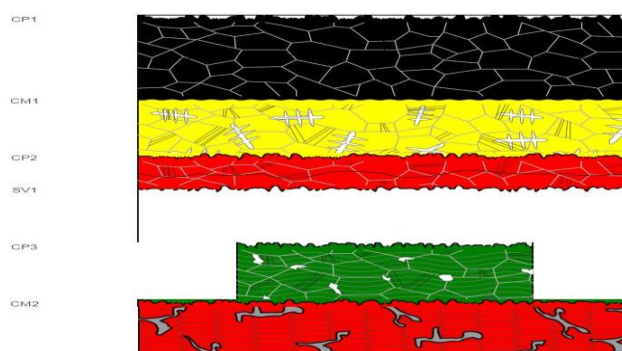


Figure 6. A visualization of stratigraphy #10 using our new tool

This prototype is not yet available for the researchers in conservation-restoration; it still needs improvements to fulfil the needs of the end users. That is why we are closely collaborating with them. Nevertheless, it provides a good idea of what the final product would look like. Visually speaking, we tried to replicate the conceptual models that conservators draw. In addition to that, we translated them so that they could be stored and processed by a computer.

In the coming months, the idea is to make this tool available through a website. That way, people working in the field of conservation-restoration will be able to use it and to supply the database with

more data. In the long run, the website should act as a platform which will attract many researchers and students. Then, the latter could consult the database and/or contribute by adding their own stratigraphies.

Conclusion

In our paper, we have talked about the steps needed in order to design data structures starting from a reality that we want to model. In our case, we already had the conceptual model which we had to translate into data structures and then into visualization of these data structures. We realized that the steps we followed were similar to what is done in many other applications.

The most critical point was the translation of the conceptual model into data structures. Indeed, as there was some information that we could not interpret by ourselves and as we needed more information than the ones that are embedded in the conceptual model, we had to collaborate closely with the researchers of the conservation-restoration faculty. This allowed us to supercharge the conceptual model and to come up with relevant data structures that could be further used.

From these data structures, we are now developing a tool that enables the possibility of adding and comparing stratigraphies between them. This would be the first tool available on the market for such an application. Even though the conservators have not started to use it yet, from what we have sounded out, they are determined to do so as soon as the tool is ready.

Eventually, the goal of this project is to facilitate the conservators' work and to provide them with an innovative tool to build their stratigraphies. In addition, this tool is a great way for the methodology developed by Bertholon to be known and used all around the world, as it replicates his stratigraphy model. When the tool is ready, we will have more feedback coming from the field of conservation-restoration. After that, we hope that a community of researchers and students will build up around our tool. Further refinements will then have to be performed, based on their comments.

Acknowledgments

This paper is partly based upon work supported by the RCSO ISnet under Award No. 38996. We would like to thank our colleagues, Christian Degrigny, Régis Bertholon and Romain Jeanneret from the Conservation-Restoration department of the Haute École Arc, who supported our work on this project with their continuous feedbacks on how to improve our decision tool to better fit their needs.

References

- Batra, S., & Tyagi, C. (2012). Comparative analysis of relational and graph databases. *IJSCE International Journal of Soft Computing and Engineering*, 2(2), 509–512.
- Bertholon, R. (2000, December 20). *La limite de la surface d'origine des objets métalliques archéologiques. Caractérisation, localisation et approche des mécanismes de conservation*. Université Panthéon-Sorbonne – Paris I.
- Cellier, F. E. & Clauss, Christoph, U. A. (2007). Electronic circuit modeling and simulation in Modelica. In *Proceedings of the Sixth Eurosim Congress on Modelling and Simulation* (pp. 1–10).
- Cypher Query Language. (2015). Retrieved June 15, 2015, from <http://neo4j.com/docs/stable/cypher-query-lang.html>
- Degrigny, C. & Senn, M. (2012). *MIFAC-Métal: Methodology to Study and Analyse the Microstructures and Corrosion Forms of Ancient and Historic Metals: Application to Metallographic Samples from Swiss Collections*. Neuchâtel.

- Holzschuher, F., & Peinl, R. (2013). Performance of graph query languages. In *Proceedings of the Joint EDBT/ICDT 2013 Workshops on - EDBT '13* (p. 195). New York, New York, USA: ACM Press.
- Neo4j. (2012). neo4j: World's Leading Graph Database. Retrieved from <http://neo4j.org/>
- Osterwalder, A., Pigneur, Y. & Tucci, C. L. (2005). Clarifying Business Models: Origins, Present, and Future of the Concept. *Communications of the Association for Information Systems*, 16, 1–25.
- Ouyang, C., Dumas, M., Aalst, W. M. P. Van Der, Hofstede, A. H. M. Ter, & Mendling, J. (2009). From business process models to process-oriented software systems. *ACM Transactions on Software Engineering and Methodology*, 19(1), 1–37.
- Peuquet, D. J. & Marble, D. F. (1990). *Introductory Readings In Geographic Information Systems*. Taylor & Francis.
- Pidd, M. (2010). *Tools for thinking: modelling in management science* (3rd Edition). Chichester, West Sussex, England: John Wiley & Sons, Inc.
- Robinson, S. (2010). Conceptual Modelling: Who Needs It? *SCS M&S Magazine*, 1–7.
- Karsten Schmidt, C. S. (2004). A Petri net Semantic for BPEL4WS - Validation and Application. In *Workshop on Algorithms and Tools for Petri Nets* (p. 6). Paderborn: Universität Paderborn.
- Wand, Y., Monarchi, D., Parsons, J., & Woo, C. (1995). Theoretical foundations for conceptual modelling in information systems development. *Decision Support Systems*, 15(4), 285–304.
- Wand, Y. & Weber, R. (2002). Research Commentary: Information Systems and Conceptual Modeling? A Research Agenda. *Information Systems Research*, 13(4), 363–376.
- White, S. A. (2004). Introduction to BPMN. *IBM Cooperation*, 2(0), 1–11.
- White, S. A. (2005). Using BPMN to Model a BPEL Process. *BPTrends*, 3(3), 1–18.

SMARTPHONE MEDICAL APP USE: A SURVEY AMONG MEDICAL STUDENTS AT ARISTOTLE UNIVERSITY OF THESSALONIKI

Complete Research

Chatzipavlou, Ioannis, University of Macedonia, Thessaloniki, Greece, gcha@uom.edu.gr

Misirlis, Nikolaos, University of Macedonia, Thessaloniki, Greece,

nikosmisirlis@uom.edu.gr

Vlachopoulou, Maro, University of Macedonia, Thessaloniki, Greece, mavla@uom.edu.gr

Abstract

Smartphones evolve in a potential learning companion among students. The specific needs of medical education have brought in light, apps dedicated to medical school. Therefore, the impact that arises in academia attracted the focus of scientific world for further research. The aim of this paper is to identify the usage of medical apps among medical students, exploring their usage patterns in order to enhance their educational activities. For that purpose, 381 students were asked to fulfill a 16-item questionnaire. The respondents (300 students) constitute a 78.7% response rate, with a Cronbach's $\alpha = 0.95$. According to the survey, 93.6% (n: 281/300) of medical students own smartphones, 64.8% (n: 182/281) of them use Android devices and 17.1% (n: 48/281) iOS devices.

The majority of students, that is 57.7% (n: 173/300) have downloaded 1 to 5 medical apps, whereas the number of students with more than 10 apps is significant lower. There is noteworthy to mention that a meaningful segment uses rarely or never the installed medical apps (59.3% and 36.7% respectively). The future development and use of medical apps to support both education and clinical practice, seems to attract the majority of the respondents.

Keywords: Smartphones, App usage, medical students, survey.

1 Introduction

Through the last decades education and technology are strictly combined, as education, the cornerstone of academic related professions, has to follow the technology imperatives and technology constitutes a means for education. Therefore, scientists have defined the modern era as “information age”, “computer age” (Abu-faraj Ziad O, Chaleby, M. H., Barakat, S.S., Zaklit, J.D. 2011) or “digital age”. The constant evolution of medical science in general, and especially of medical services, has created the necessity of innovations and alterations in medical education. The medical science adopted the new status quo in all its facets and this inevitably assimilated to education.

The debut of Personal Digital Assistants (PDAs) helped medical students to find a useful companion in their campus lives (Lindquist, Johansson, Petersson, Saveman, & Nilsson, 2008), (Economides & Nikolaou, 2006). Particularly, they utilize the device’s high portability, small size and large computational power to assess them in classroom for formal instruction (Kho, Henderson, Dressler, & Kripalani, 2006) as well as in clinics (Littman-quinn et al., 2011). In 2007, the scenery changed more by the introduction of Apple’s smartphone, iPhone (Wallace, Clark, & White, 2012). The large, high sensitivity touchscreen (Hakobyan, Lumsden, O’Sullivan, & Bartlett, 2013), the high performance processors (Gutierrez et al., 2011), the ease of use (Franko & Tirrell, 2012) and the massive increase of mobile internet bandwidth are the key components that made smartphones omnipresent in academia (Vinay & Vishal, 2013). Tablets and smartwatches, as well, endow medical students with innovative means of education. In addition to the aforementioned advantages, smartphones are equipped with microsensors and a high resolution camera, possible to collect patients’ data (Vinay & Vishal, 2013). These last additions boost the use of such tools among students, professors and patients, especially with the advent of smartphone initiating the use of apps.

Recent studies (Robinson et al., 2013), (Payne, Wharrad, Watts, & Payne, 2012)(Rung, Warnke, &Mattheos, 2014)show a continuously increase of medical app usage among students. This distinct group of students is described by singularities, such as preclinical and clinical curricula. In this point of view, studying apps’ usage among medical students differs significantly across other students. Understanding all the aspects around apps, smartphones and academic life regarding medical students, is therefore crucial. The contribution of this paper is to examine the use of smartphones among medical students as an m-learning tool. Furthermore, the results distinguish between the most preferred apps taking into consideration the frequency of their usage. Finally, the paper contributes by firstly reporting the current situation on the field on Greek universities.

The paper is organized as follows: Section 2 presents the methodology that followed to design and distribute the questionnaires; Section 3 describes the results; Section 4 discusses the results; Section 5 adds limitation to the research; finally in Section 6 concludes the paper and presents future work that can be applied.

2 Methodology

Firstly, it is necessary to make a note regarding the specific medical school, in order to clarify the particularities and the differences from other relevant schools, globally. In Greek medical schools, there is no specific separation between preclinical and clinical courses in relation with the years of study. Precisely, at the first five semesters, students attend both theoretical and laboratory courses. From the sixth semester and thereafter, clinical courses are incorporated to the syllabus as well. Specifically, students must attend theoretical and clinical courses as well as laboratories. Though, during these seven semesters, laboratory courses are continuously decreasing, but clinical courses are increasing. Finally, to clarify more the Greek medical school system, it can be noticed that each clinical course is accompanied by the equivalent theoretical course. Hereinafter the theoretical and laboratory courses will be called Education and the rest Clinical in order to facilitate the reading.

We designed a 16-items (both multiple choice and free response question) questionnaire which was distributed on person and the answers were collected manually. We collected data from 300 medical undergraduate students of Aristotle's University of Thessaloniki in Greece. 81 students left the questionnaire unfilled, achieving a response rate equal to 78.7%. The high response rate is due to the manually procedure of collecting data we followed. The Aristotle University is the biggest academic institute in Greece with more than 90000 active students and among them 4000 medical students. Our sample represents a 7.5% of medical students in Thessaloniki. The answers derived from students of all academic year range.

Our survey was based on previous literature (Garritty & El Emam, 2006), (Payne et al., 2012), (Robinson et al., 2013), (Rung et al., 2014). Before the distribution, the questionnaire was revised by physicians and consequently was tested on a short number of medical undergraduate students in order to check the content's comprehensiveness, validity and reliability.

The questionnaire was structured in four sessions: Demographic and mobile device ownership, installed medical apps, medical apps usage and medical app proposed improvement.

The first session, encapsulates, except the demographic, the main mobile device used by students, as well as its operation system (OS). The second session consists on the assessment of certain app use. Specifically, we asked students which type of apps they use related to their specific objectives. The students, in the

third session, were called to answer about the frequency of app usage. Furthermore, they responded about the specific app use, on the exercise of their medical duties. The questionnaire concludes with questions about medical app proposed improvement.

The collected numerical data were entered, processed and analyzed with the statistical software SPSS (version 22). For the evaluation of our findings and the statistical significance we conducted inferential analysis (Chi square and linear regression).

3 Results

We first tested the 16 items on their internal consistency. Using reliability analysis we obtained a Cronbach's Alpha of 0.95, indicating a very acceptable fit.

3.1 Demographic characteristics

In total 300 medical students answered the questionnaire giving a response rate equal to 78.7%. Of these, 144 (48%) are males and 156 (52%) are females. From a total of 300 students, 26 (8.67%) are first-year students, 14 (4.67%) second-year, 69 (23.00%) third, 117 (39.33%) fourth, 28 (9.33%) fifth and 45 (15.00%) are sixth-year students of the medical school.

3.2 Type of mobile devices and operating systems

A total of 1(0.3%) student owns feature phone (Schiefer& Decker, 2005), 281 (93.6%) use smartphones, and 18 (6.0%) use both smartphones and tablets.

Android is the main and most common operating system used by 182 users (64.8%), and iOS with 17.1% and 48 users, the second most used. Symbian is used by the 11.00% and 33 users.

3.3 Installed medical Apps

The highest percentage of students own from 1 to 5 medical applications. On the other hand, 33.7% of the students do not own any medical application (Table 1).

Table1: Number and percentage of health applications

| Number of medical applications | Percentage of use among users (number of users/ total users) |
|---------------------------------------|---|
| None | 33.7% |
| 1 – 5 | 57.7% |
| 6 – 10 | 6.3% |
| 11 – 15 | 1.0% |
| 16+ | 1.3% |

Students were asked, further, if fee is an inhibitory factor for the app acquisition. Observing the results, 2% of the students paid to purchase a medical app against 87.9% who preferred not to.

3.4 Medical Apps Usage

Among those who have at least one medical app, 159 responded that they use these apps to support their education on medical school and 114 use the apps in clinics. Table 2 represents the percentages of app usage per year of study. We divide the categories of usage in two segments: education and clinical.

Further on, students were asked about the frequency of usage. Table 3 collects these results dividing apps in education and clinical, as on Table 2.

Furthermore, we asked, which app is used more frequently during the courses, education and clinical. Results show that students use: 60.8% (n: 171/281) informative medical app for education support, 34.5% (n: 97/281) diagnosis-related apps, 27.4% (n: 77/281) first-aid apps, 19.6% (n: 55/281) clinical score apps (bpm etc.), 8.9% (n: 25/281) health record management apps, and finally 1.8% (n: 5/281) other medical apps.

Next figure depicts the frequency of use, correlated with the different medical app categories.

Table 1: Medical app usage / Year of study

| Year of study | Education (number of students/ number of total students per specific year) | Clinical (number of students/ number of total students per specific year) |
|----------------------|---|--|
| 1 st | 50%(13) | 35% (9) |
| 2 nd | 28% (4) | 33% (3) |
| 3 rd | 52% (36) | 33% (23) |
| 4 th | 58% (68) | 41% (48) |
| 5 th | 61% (17) | 50% (14) |
| 6 th | 47% (21) | 38% (17) |

Table 2: Medical app frequency of use

| Frequency | Education (number of students/ number of total students per specific year) | Clinical (number of students/ number of total students per specific year) |
|------------------|---|--|
| Many times/day | 17.1% (34) | 8.6% (17) |
| 1-2 times/day | 14.6% (29) | 9.0% (18) |
| Many times/week | 20.1% (40) | 9.0% (18) |
| 1-2 times/week | 14% (28) | 11.6% (23) |
| Rarely | 19.1% (38) | 40.2% (80) |
| Never | 15.1% (30) | 21.6% (43) |

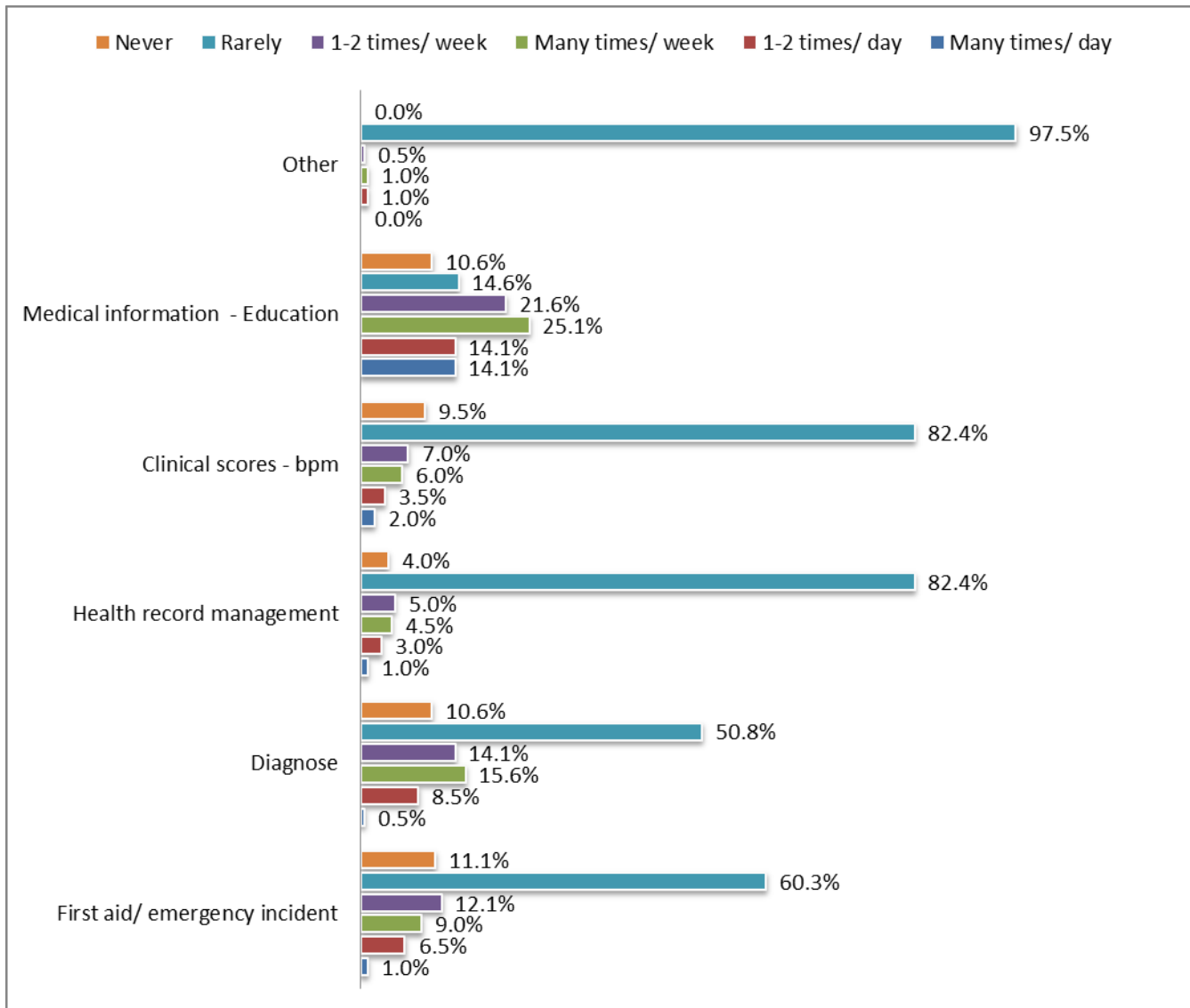


Figure 1: Medical app category - frequency of use

3.5 Future App Design

Lastly we have asked students if they have an opportunity to design an app in which specific field might it be. The majority chooses apps related to medical information and education in a percentage of 33.7% (n: 101/300). In continuous, 11.7% (n: 35/300) answer clinical scores, 15.3% (n: 46/300) health record management, 21.3% (n: 64/300) diagnose, 13.3% (n: 40/300) emergency incidents and 4.7% (n: 14/300) other categories.

4 Discussion of results

To our knowledge, this is the first study to examine examines smartphone ownership and usage between medical students in Greece. The collected data could be a useful component to form smartphone ownership and usage landscape. More general, it can be observed a European tendency on health apps promotion. This can be extended in medical education rendering our research timely (European commission, Brussels, Green paper on mobile Health).

4.1 Smartphone ownership

Previous research by Payne et al. (2012) finds smartphone ownership equal to 79% and shows an equal preclinical and clinical distribution. Our study finds a 96.6% ownership percentage. This contrast of 17.6% can be explained by the increased penetration of smartphones in worldwide population that occurred in the two year period between Payne’s and present research.

In our case preclinical and clinical students possess smartphones by 100% both (n1: 40/40, n2: 190/190). Because of the separation that occurs in fifth semester, it is impossible to identify which students of the third year belong to which semester (preclinical and clinical. This is why the third year results about smartphone ownership are excluded from being summarized.

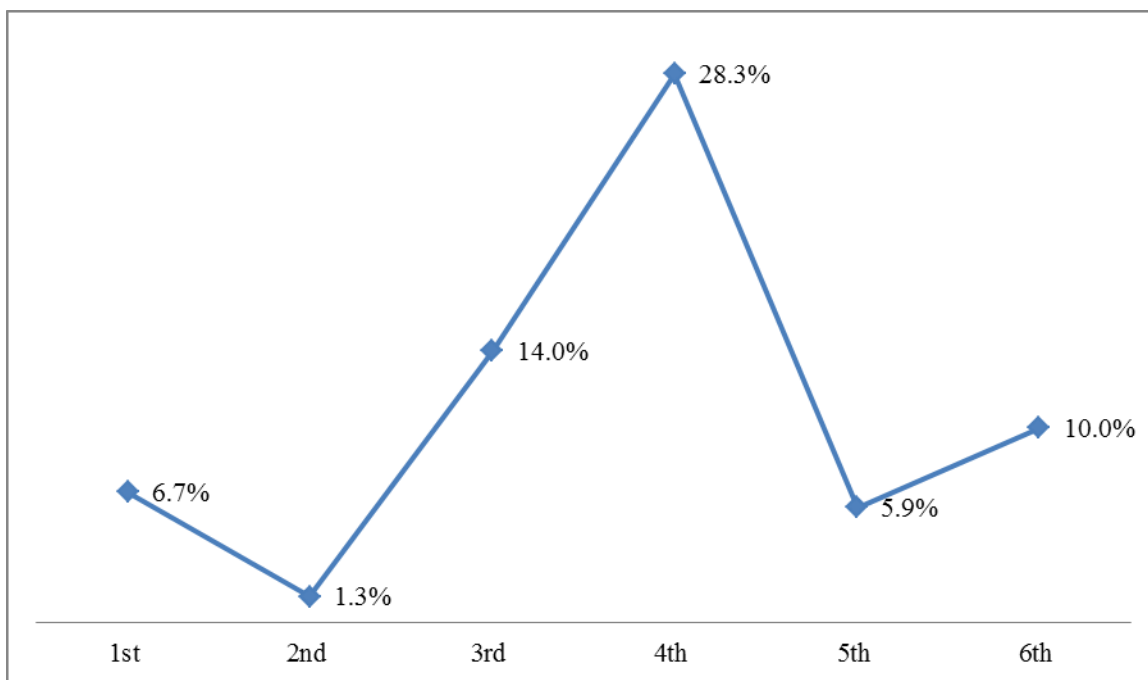


Figure 2: Percentage of installed apps/ year of study

4.2 Number, frequency and type of medical apps

As far as concern, the amount of medical apps that a student has installed, the segment that prevails is *1 to 5 apps* with 57.7% (Table 1). Can be easily noticed (Figure 2) that during the studies, students at fourth year install more apps. This trend starts in the fifth semester, in which clinical courses begin; students tend to experiment in diverse medical apps. In addition, most of the students do not pay to purchase an application but prefer to download free or freemium apps. The aforementioned observations can lead us to the sequent outcome: student lose their interest about medical apps. This statement can be affirmed also from figure 1 where it shows that the majority of the students, 60% and more, have installed apps but use them rarely or never. The only exception is the category *Medical information – Education*, in which the percentage of *Never* and *Rarely* is 25.2%. Expanding this thought we can say that apps in all other categories need a bigger how-to knowledge, so the student installs but quickly abandons, tending to use only apps which are easier to master.

This derives from the lack of proper app consulting by the university personnel, following by a plethora of unregulated (Murfin, 2013) and low quality apps, available on app marketplaces. Lastly, from table 2 we conduct the following conclusion: students, even though in clinical years, they do not lose their interest in education apps.

Another interest outcome arises from the number of the surveyed students who use Android or iOS smartphones associated with the type of app installed, education or clinical. Precisely, 68.6% (n: 114/182) and 48.3% (n: 88/182) of Android users hold education and clinical apps respectively. On the other hand the relative percentages of iOS users are: 18.8% (n: 9/48) and 6.3% (n: 3/48). It is unclear if these significant differences are due to absence of iOS apps and/ or iOS appstore policy.

4.3 Designing a medical app

From students' answers it is clear that the majority concerns about apps related to medical information and education in a percentage of 33.7% (n: 101/300). In combination with figure 2 we notice that students want to design apps that use more frequent. The finding suggests that students can recognize weaknesses of an app by using them during their education and therefore recommend improvements.

5 Limitations

The main concern of the paper was to present data about smartphone ownership and usage among medical students. This study has several limitations. Firstly, there are no questions about postgraduate students, junior doctors or medical professors. In continuous, our study limited to only one university

students. Another limitation of the research was that medical students have a lack of experience in medical app use and tend to experiment in this field. A possible outcome of this behavior is that they can answer in an unpredictable manner. Lastly there was a difficulty to adjust the questionnaire to the Greek medical school curriculum. More specific the segmentation between preclinical and clinical years of study is unclear.

6 Conclusions

The medical profession is undoubtedly correlated with the usage of smartphone technology (Franko & Tirrell, 2012). Particularly there is a constant growth of medical applications globally. Our research focus on medical students in Greece and shows that local reality is far from that standard. This deduction derives from the percentage of students that own smartphone but do not use at all or rarely use medical apps. This amount of students features a lack of usage, penetration and adoption of smartphone technology devices. The previously mentioned conclusion is explained by the skepticism among users in regard of existing medical apps reliability. This outcome is further clarified from the fact that students were asked if an app designed by them would be useful for educational purposes. The majority of them answered positively.

This can reveal opportunities for Greek academia to adopt formally smartphones and consequently medical apps to educational process. In advance new medical apps can be designed to assist furthermore the academic procedure.

References

- Abu-faraj Ziad O, Chaleby, M. H., Barakat, S.S., Zaklit, J.D. (2011). A SIM Card-Based Ubiquitous Medical Record Bracelet / Pendant System - A Pilot Study. *Science And Technology*, 1914–1918.
- Economides, A. a, & Nikolaou, N. (2006). Evaluation of Handheld Devices for Mobile Learning. *International Journal of Engineering Education*, 24(1), 1–21. Retrieved from <http://www.conta.uom.gr/conta/publications/PDF/Evaluation of Handheld Devices for Mobile Learning.pdf>
- Franko, O. I., & Tirrell, T. F. (2012). Smartphone app use among medical providers in ACGME training programs. *Journal of Medical Systems*, 36(5), 3135–9. <http://doi.org/10.1007/s10916-011-9798-7>
- Garrity, C., & El Emam, K. (2006). Who's using PDAs? Estimates of PDA use by health care providers: A systematic review of surveys. *Journal of Medical Internet Research*, 8(2). <http://doi.org/10.2196/jmir.8.2.e7>
- Gutierrez, A., Dreslinski, R. G., Wenisch, T. F., Mudge, T., Saidi, A., Emmons, C., ... Austin, A. R. M. (2011). Full-System Analysis and Characterization of Interactive Smartphone Applications. *Architecture*, 81–90.
- Hakobyan, L., Lumsden, J., O'Sullivan, D., & Bartlett, H. (2013). Mobile assistive technologies for the visually impaired. *Survey of Ophthalmology*, 58(6), 513–28. <http://doi.org/10.1016/j.survophthal.2012.10.004>
- Kho, A., Henderson, L. E., Dressler, D. D., & Kripalani, S. (2006). Use of handheld computers in medical education. A systematic review. *Journal of General Internal Medicine*, 21(5), 531–7. <http://doi.org/10.1111/j.1525-1497.2006.00444.x>
- Lindquist, A. M., Johansson, P. E., Petersson, G. I., Saveman, B.-I., & Nilsson, G. C. (2008). The use of the Personal Digital Assistant (PDA) among personnel and students in health care: a review. *Journal of Medical Internet Research*, 10, e31–e31. <http://doi.org/10.2196/jmir.1038>
- Littman-quinn, R., Chandra, A., Schwartz, A., Chang, A. Y., Fadlemola, F. M., Ghose, S., ... Bewlay, L. (2011). mHealth Applications for Clinical Education , Decision Making , and Patient Adherence in Botswana, 1–8.
- Murfin, M. (2013). Know your apps: an evidence-based approach to evaluation of mobile clinical applications. *The Journal of Physician Assistant Education : The Official Journal of the Physician Assistant Education Association*, 24(3), 38–40. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/24261171>
- Payne, K. F. B., Wharrad, H., Watts, K., & Payne, K. F. B. (2012). Smartphone and medical related App use among medical students and junior doctors in the United Kingdom (UK): a regional survey. *BMC Medical Informatics and Decision Making*, 12, 121–121. <http://doi.org/10.1186/1472-6947-12-121>
- Robinson, T., Cronin, T., Ibrahim, H., Jinks, M., Molitor, T., Newman, J., & Shapiro, J. (2013). Smartphone use and acceptability among clinical medical students: A questionnaire-based study. *Journal of Medical Systems*, 37(3). <http://doi.org/10.1007/s10916-013-9936-5>
- Rung, A., Warnke, F., & Mattheos, N. (2014). Investigating the use of smartphones for learning purposes by Australian dental students. *JMIR mHealth and uHealth*, 2(2), e20–e20. <http://doi.org/10.2196/mhealth.3120>
- Schiefer, G., & Decker, M. (2005). Taxonomy for mobile terminals, 255–258.
- Vinay, K. V., & Vishal, K. (2013). Smartphone applications for medical students and professionals, 3(1), 59–62.

Wallace, S., Clark, M., & White, J. (2012). 'It's on my iPhone': attitudes to the use of mobile computing devices in medical education, a mixed-methods study. *BMJ Open*, 2(4), 1–7. <http://doi.org/10.1136/bmjopen-2012-001099>

THE DETERMINANTS OF CLOUD COMPUTING ADOPTION BY LARGE EUROPEAN FIRMS

Research in Progress

Xiaolin Cheng, University of Paris-Sud, Sceaux, France, xiao-lin.cheng@u-psud.fr
Ahmed Bounfour, University of Paris-Sud, Sceaux, France, ahmed.bounfour@u-psud.fr

Abstract

This paper reports ongoing research into the human and non-human factors underlying the adoption of cloud computing in large European firms. To assess these determinants, we develop a model based on the Technology Acceptance Model (TAM) and the Technology-Organization-Environment (TOE) framework. We develop a questionnaire to collect data on each of variables identified in the proposed model.

Keywords: *Cloud computing adoption, TOE and TAM frameworks, large European firms, Survey methods, Regression analysis*

Introduction

Theoretically, cloud computing offers a vast opportunity for organizations and enterprises to improve the flexibility and smooth operation of their business models (Accenture, 2012). Consequently, it would appear that these organizations would find it very easy to migrate to cloud computing. However, in practice, the debate rages regarding cloud adoption (Khajeh-Hosseini et al., 2010). As cloud computing requires a high level of commitment, research frameworks may help to reduce the perceived technical, financial and social risks (Barki, 2007; Eze et al., 2011; Silva, 2007).

Early studies on cloud computing adoption tend to be skewed toward benefits and challenges (Lorraine, M. et al. 2013), and only a few have addressed it from a human perspective (see Table 1). There is currently no holistic evaluation of the human and non-human determinants of adoption. In this study we address this crucial gap by developing an integrative research model that combines the Technology Acceptance Model (TAM) and Technology-Organization-Environment (TOE) framework. We use the model to evaluate the determinants of cloud computing adoption in an organization. This study therefore presents a more holistic assessment of the determinants of cloud computing adoption than earlier work. In addition, this investigation of the determinants of adoption in large European firms contributes to the wider body of scientific knowledge that has so far paid little attention to the issue.

In the domain of Information System (IS), theories aim to understand, explain, or predict how, why and to what extent individuals, firms or organizations adopt and agree to deploy a new technology. Based on this earlier work, which forms the background for the problem, the main research question for this study is: what are the key determinants that influence the organizational adoption of cloud computing by large European firms?

The following sub-questions are framed to address the research problem.

1. What are the theoretical frameworks for the adoption of cloud computing by large European firms?
2. How does the heterogeneous network of human and non-human factors influence the adoption of cloud computing by large European firms?

Literature review

Cloud computing research is still in its early days. Much of the current literature focuses on its benefits and risks, organizational case studies of cloud adoption and cloud computing architectures (Bhattacharjee and Park, 2013). A methodological approach to the comparison of cloud versus in-house solutions has been proposed that is based on an assessment of the direct economic impact of migration to the cloud (Naldi and Mastroeni, 2014). This approach uses Net Present Value and stochastic models of storage costs and memory requirements. The adoption of new technologies, such as cloud computing, is a complex phenomenon with a high level of ambiguity and a variety of opportunities and challenges. We carried out a detailed analysis of the literature on cloud computing adoption based on the articles summarized in Table 1.

Table 1 Literature review

| Article | Framework | Constructs | | | |
|---------------------------|------------------|------------|--------------|-------------|-------|
| | | Technology | Organization | Environment | Human |
| (Lin & Chen, 2012) | DOI | √ | | | |
| (Ratten, 2012) | Social cognitive | | | | √ |
| (W. W. Wu, 2011) | TAM | √ | | | √ |
| (Lee et al., 2013) | PEST | √ | | √ | |
| (Misra & Mondal, 2011) | No framework | | | √ | |
| (Low, Chen & Wu, 2011) | TOE | √ | √ | √ | |
| (Gupta et al., 2013) | No framework | √ | | | |
| (Hsu et al., 2014) | TOE | √ | √ | √ | |
| (Nkhoma & Dang, 2013) | TOE | √ | √ | √ | |
| (Che Hussin et al., 2013) | TOE | √ | √ | √ | |
| (Alshamaila et al., 2013) | TOE | √ | √ | √ | |
| (Lee, Park & Lim, 2013) | PEST | √ | | √ | |
| (Wu et al., 2013) | DOI | √ | | | |
| (Oliveira et al., 2014) | TOE and DOI | √ | √ | √ | |
| (Borgman et al., 2013) | TOE | √ | √ | √ | |
| (Benlian & Hess, 2011) | No framework | √ | | | |
| (Cegielski et al., 2012) | TOE | √ | √ | √ | |
| (Current study) | TOE and TAM | √ | √ | √ | √ |

Note: Diffusion of Innovation (DOI); Technology Acceptance Model (TAM); Technology-Organization-Environment (TOE); Political-Economic-Social and Technological (PEST)

This literature review helped us to understand the background for cloud computing adoption and related research. We then developed a research model that highlights nine factors influencing cloud adoption. Subsequently, we created a questionnaire in order to collect empirical data and test our hypotheses. In further research, the data will be analyzed using the Partial Least Square (PLS) method. The results of this study will not be only useful for large European firms who must decide whether to adopt the new technology, but also as a basis for further research into this subject.

Research model and hypotheses

The Technology Acceptance Model (TAM) (Davis, 1989) was the first attempt to develop an overall approach to the issue of adoption in the domain of IS (Barki, 2007; Eze et al., 2011; Silva, 2007). The model considers perceived usefulness (PU) and perceived ease of use (PEOU) as the key determinants of the adoption of information technology. Despite its widespread diffusion and implementation in IS research, the model suffers from the narrow focus on only two main dimensions, while other use factors are ignored. Other research has extended the scope of the analysis and added other dimensions: the Technology, Organization and Environment (TOE) framework is a notable example (Awa, H.O. & Vkoah, O. 2012).

The TOE framework identifies various influential factors in the innovation adoption process (Tornatzky & Klein, 1982). There are features of the TOE framework that make it appropriate for the investigation of cloud computing adoption, which is unlike conventional innovation adoption and diffusion scenarios. Cloud computing services are usually provided to firms and organizations by a third party (cloud service providers). Thus, unlike conventional innovations, cloud computing technology has three main players: cloud-based services, cloud users and cloud service providers. As a result, its adoption is influenced by three major factors: 1. the characteristics of cloud computing technology, which is a function of both

technologies that are both internal and external to the company; 2. the characteristics and resources of firms and organizations that provide the organizational context; 3. the characteristics of the environmental context in which a firm conducts its business: its industry; competitors; access to resources supplied by others; and dealings with the government.

In this study, we integrate constructs from both the TOE method and the TAM framework in order to include both human and non-human actors in the network. Based on the references given in Table 2, we developed the hypotheses outlined below.

Table 2 *Constructs and their resources*

| Constructs | Variables | References |
|--------------|------------------------|--|
| Human | Perceived usefulness | (Oliveira et al., 2014), (Moore & Benbasat, 1991) |
| | Perceived ease of use | (Moore & Benbasat, 1991), (Gupta et al., 2013) |
| Technology | Complexity | (Low et al., 2011), (Moore & Benbasat, 1991), (Ifinedo, 2011) |
| | Compatibility | (Low et al., 2011), (To & Ngai, 2006), (Ifinedo, 2011), (Zhu, Dong, Xu & Hally, 2006) |
| Organization | Top management support | (Low et al., 2011), (Zhu, Li, Wang & Chen, 2010), (Chwelos, Benbasat & Dexter, 2001) |
| Environment | Adequate resources | (Chang et al., 2007) |
| | Vendor support | (Thong, 2001) |
| | Government policy | (Zhu & Kraemer, 2005), (Marston et al., 2011) |
| | Competitive pressure | (Low et al., 2011), (To & Ngai, 2006), (Pan & Jang, 2008), (Ifinedo, 2011) |

Hypotheses

- **Perceived usefulness**

Perceived usefulness refers to the degree to which a person believes that the use of a system will improve his performance.

H1 Perceived usefulness is positively related to organizational adoption of cloud computing

- **Perceived ease of use**

Perceived ease of use is described as the degree to which a person believes that the use of a system will be effortless.

H2 Perceived ease of use is positively related to organizational adoption of cloud computing

- **Complexity**

Complexity describes “the degree to which an innovation is perceived as difficulty to understand and to use”

H3 Complexity is negatively related to organizational adoption of cloud computing

- **Compatibility**

Compatibility reflects the degree to which an innovation is perceived as being consistent with the existing values, past experiences, and needs of receivers.

H4 Compatibility is positively related to organizational adoption of cloud computing

- **Top management support**

Top management support can contribute to the adoption of innovation by creating a fertile environment and providing resources (Premkumar & Roberts, 1999).

H5 Top management support is positively related to organizational adoption of cloud computing

- **Adequate resources**

Adequate resources are also critical to the success of adoption. If the budget is sufficient, positive support will be provided by human resources.

H6 Adequate resources are positively related to organizational adoption of cloud computing.

- **Vendor support**

In cloud computing technology, the customer is highly dependent on the vendor to achieve the desired level of security. This dependency is highlighted in low-tech companies, which lack information technology expertise. Therefore, vendors must guarantee security, availability, and performance through clear Service Level Agreements (SLA)

H7 Vendor support is positively related to organizational adoption of cloud computing

- **Government policy**

Government policy is another environmental factor that affects innovation diffusion. Companies operating in an environment where government policies are restrictive have a low level of technology adoption.

H8 Government policy is negatively related to organizational adoption of cloud computing

- **Competitive pressure**

Competition pressure is defined as “the degree that the company is affected by competitors in the market” (Zhu, Xu & Dedrick, 2003). Firms react by adjusting their offers and greater competition forces a firm to allocate more resources to innovation.

H9 Competitive pressure is positively related to organizational adoption of cloud computing.

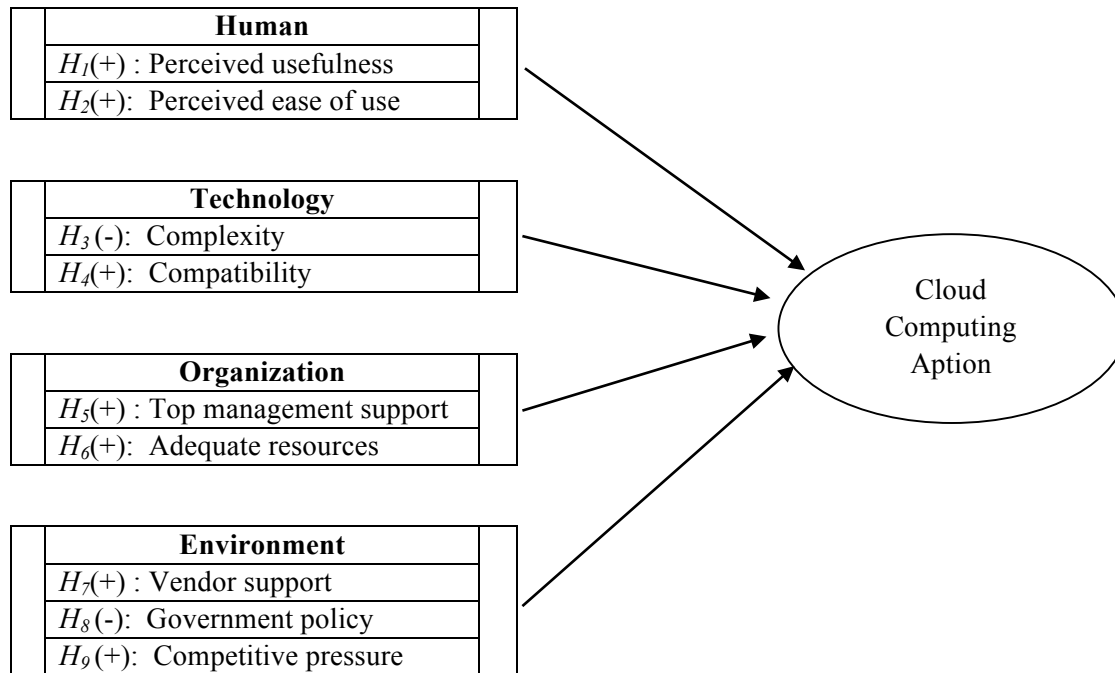


Figure 1 Research Model

Method

To test the hypotheses outlined above, we used a questionnaire to collect data on each of the variables given in the proposed model. The sample consisted of large European firms in high-tech industries. Variables were measured on a five-point Likert scale, with values ranging from 1 (strongly disagree) to 5 (strongly agree). Information technology (IT) staff and managers of high-tech firms were targeted as they are best placed to understand their firm’s current IT operations and future trends.

The first step consisted of a pilot study designed to ensure that the questionnaire was suitable for the research context. This consisted of personal interviews that aimed to verify the efficiency of the questionnaire. Associate professors of IT (specialized in cloud computing) and other IT professionals were asked to review the questionnaire. Based on their feedback, some items were modified to improve clarity. The questionnaire itself will be distributed using the firms’ own platforms, to increase the response rate. The constructs in the questionnaire were developed from the literature reviews described in Table 1 and Table 2, and were modified to fit the context of cloud computing adoption. The questionnaire was divided into two parts:

1. Demographic characteristics including: the number of employees; capital; annual revenue; and adoption of cloud computing services
2. Evaluation of the nine predictors using a five-point Likert scale

Table 3 was constructed to support the hypotheses, from which we can find the loadings for each hypothesis. And all the used references were outlined in Table 2.

Table 3 Measurements and their loadings

| Measurements | Loadings |
|------------------------|---|
| Perceived usefulness | <p>Compared to current technologies, cloud computing enables tasks to be accomplished more efficiently</p> <p>Cloud computing contributes to the agility of the enterprise</p> <p>Cloud computing technology helps to reduce operational maintenance, updating and training costs</p> |
| Perceived ease of use | <p>A good internet connection and speed of cloud services are necessary</p> <p>Cloud tools and data can be used and accessed from anywhere</p> <p>Negligible learning time for all employees</p> |
| Complexity | <p>Cloud computing is too complex for business operations</p> <p>The skills needed to adopt cloud computing are too complex for employees</p> <p>The complexity of transferring current systems to a cloud computing platform is a concern</p> <p>Uncertainty about the location of data limits the use of cloud computing services</p> <p>The risk of a security breach limits the use of cloud computing services</p> |
| Compatibility | <p>Cloud computing technology is compatible with current practices</p> <p>Cloud computing technology is consistent with the firms' core values and goals</p> |
| Top management support | <p>Cloud technology can easily be integrated into existing IT infrastructure</p> <p>Applications are loosely coupled and independent</p> <p>The company's top management provides strong leadership and engages in information systems' policy</p> <p>The company's top management is willing to take risks in the adoption of cloud computing</p> |
| Adequate resources | <p>Does the firm have enough resources to support the development of cloud computing technology?</p> <p>Does the firm have enough time to develop cloud computing technology?</p> <p>Does the firm have an adequate budget for the development of cloud computing technology?</p> <p>Does the firm have enough human resources to develop cloud computing technology?</p> |

| | |
|----------------------|--|
| Vendor support | <p>There is a service level agreement (SLA), guaranteed by the vendor</p> <p>The vendor will cooperate in returning data if the company decides to change vendor</p> <p>The firm will receive adequate compensation for a vendor breach of the SLA</p> <p>Support is easily available from cloud computing vendors during implementation</p> <p>Adequate training is provided by vendors</p> |
| Government policy | <p>The government encourages firms to apply cloud computing</p> <p>There are intermediate organizations that support enterprises during implementation</p> <p>Appropriate regulations are in place to deal with any legal challenges related to cloud computing</p> |
| Competitive pressure | <p>The firm thinks that cloud computing has an influence on competition in their industry</p> <p>The firm is under pressure from competitors to adopt cloud computing</p> |
| Intention to adopt | <p>The firm has already planned to use cloud computing services</p> <p>The firm plans to adopt cloud computing in the next year</p> <p>Cloud computing is a solution that the firm is ready to accept</p> |

All items are based on five-point scale

Analysis

There are several analytic methodologies for measuring cloud computing adoption, such as the Mann-Whitney test (Borgman et al., 2013) and hierarchical regression (Cegielski, 2012). The Partial Least Square (PLS) approach has emerged as the most popular and widely-used methodology due to its effectiveness and ease of use.

The methodology has been used to validate measurements and to test hypotheses. We will test our research hypotheses using PLS-based structural equation modeling. Unlike parameter-oriented and covariance-based structural equation modeling, the component-based method is prediction-oriented and places minimal restrictions on sample size and residual distributions. Furthermore, there are no data normality requirements. The technique employs a component-based approach for model estimation and is best suited for testing complex structural models. A two-step approach will be used: first, the quality of measures will be assessed using the measurement model; then the hypotheses will be tested using the following structural model. Let $X = \{X_i \mid i=1, 2, \dots, 9\}$ be the set of determinants, $\beta = \{\beta_i \mid i=0, 1, \dots, 9\}$ be the parameters and Y was defined as the intention of cloud computing adoption.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \mu$$

Although we assume that all the explanatory variables are independent in this equation, in reality there are interdependencies. To address this limitation, we will model the relationships between multiple attributes in the analysis of the survey results.

Conclusion and future research

No previous research has drawn on innovation adoption theory that considers both the human and non-human factors that impact cloud computing adoption. This study addresses this gap in the research. It combines the TAM method and TOE framework in order to assess the factors influencing the adoption of cloud computing from human and non-human perspectives. A questionnaire was designed to collect data related to the research variables. Future research will focus on the development and diffusion of the questionnaire and the analysis of the results.

References

- Accenture. (2012). A new era in banking cloud computing changes the game. Accenture.
- Alshamaila, Y., Papagiannidis, S., & Li, F. (2013). Cloud computing adoption by SMEs in the north east of England: A multi-perspective framework. *Journal of Enterprise Information Management*, 26, 250–275.
- Awa, H.O., Ukoha, O. (2012). Integrating TAM and TOE frameworks and expanding their characteristic constructs for e-commerce adoption by SMEs. *Proceedings of Information Science & IT Education Conference (InSITE)*
- Barki, H. (2007). Quo vadis, TAM? *Journal of the Association Fo Information Systems*, 8(4), 211–218.
- Benlian, A., & Hess, T. (2011). Opportunities and risks of software-as-a-service: Findings from a survey of IT executives. *Decision Support Systems*, 52(1), 232–246.
- Bhattacharjee A. and Park SC. (2013). Why end-users move to the cloud: a migration-theoretic analysis. *European Journal of Information Systems* 1-16
- Borgman, H. P., Bahli, B., Heier, H., & Schewski, F. (2013). Cloudrise: Exploring cloud computing adoption and governance with the TOE framework. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 4425–4435.
- Cegielski, C. G., Jones-Farmer, L. A., Wu, Y., & Hazen, B. T. (2012). Adoption of cloud computing technologies in supply chains: An organizational information processing theory approach. *International Journal of Logistics Management*, 23(2), 184–211.
- Chang, I. C., Hwang, H. G., Hung, M. C., Lin, M. H., & Yen, D. C. (2007). Factors affecting the adoption of electronic signature: Executives' perspective of hospital information department. *Decision Support Systems*, 44, 350–359.
- Chwelos, P., Benbasat, I., & Dexter, A. S. (2001). Research Report: Empirical Test of an EDI Adoption Model. *Information Systems Research*, 12(604), 304–321.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340
- Davis, F. D. (1989). User Acceptance of Computer TEchnology: A Comparison of Two Theoretical Models. *Management Science*, 35(8), 982-1003
- Davis, F. D. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International Journal of ManMachine Studies*.
- Eze, S. C., Okoye, J. C., Nebo, O. G., Ohakwe, S. N., Chukwuemeka, E., & Anazodo, R. (2011). Using the Characteristics of Small Business Managers to Understand Information Technology (IT) Adoption in Nigeria. *International Journal of Business and Social Science*, 2(13), 82–91.
- Gupta, P., Seetharaman, a., & Raj, J. R. (2013). The usage and adoption of cloud computing by small and medium businesses. *International Journal of Information Management*, 33(5), 861–874.
- Hsu, P. F., Ray, S., & Li-Hsieh, Y. Y. (2014). Examining cloud computing adoption intention, pricing mechanism, and deployment model. *International Journal of Information Management*, 34(4), 474–488.

- Ifinedo, P. (2011). Internet / e-business technologies acceptance in Canada ' s SMEs : Focus on Organizational and Environmental Factors.
- Khajeh-Hosseini, A., Greenwood, D., Smith, J. W., & Sommerville, I. (2010). The Cloud Adoption Toolkit: Addressing the Challenges of Cloud Adoption in Enterprise.
- Lee, S. G., Chae, S. H., & Cho, K. M. (2013). Drivers and inhibitors of SaaS adoption in Korea. *International Journal of Information Management*, 33(3), 429–440.
- Lee, S., Park, S. B., & Lim, G. G. (2013). Using balanced scorecards for the evaluation of “Software-as-a-service.” *Information & Management*, 50(7), 553–561.
- Lin, A., & Chen, N.-C. (2012). Cloud computing as an innovation: Perception, attitude, and adoption. *International Journal of Information Management*, 32, 533–540.
- Lorraine, M. et al. (2013). Factors affecting the adoption of cloud computing: an exploratory study. *Proceedings of the 21st European Conference on Information Systems*.
- Low, C., Chen, Y., & Wu, M. (2011). Understanding the determinants of cloud computing adoption. *Industrial Management & Data Systems*, 111, 1006–1023.
- Lu, J., Yu, C.-S., Liu, C., & Yao, J. E. (2003). Technology acceptance model for wireless Internet. *Internet Research*, 13, 206–222.
- Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., & Ghalsasi, A. (2011). Cloud computing - The business perspective. *Decision Support Systems*, 51(1), 176–189.
- Misra, S. C., & Mondal, A. (2011). Identification of a company's suitability for the adoption of cloud computing and modelling its corresponding Return on Investment. *Mathematical and Computer Modelling*, 53(3-4), 504–521.
- Moore, G., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*.
- Naldi M. and Mastroeni L. (2014). Economic decision criteria for the migration to cloud storage. *European Journal of Information Systems*. 1-13
- Nkhoma, M. Z., & Dang, D. P. T. (2013). Contributing Factors of Cloud Computing Adoption : a Technology - Organisation - Environment Framework Approach. *International Journal of Information Systems and Engineering*, 1, 38–49.
- Oliveira, T., Thomas, M., & Espadanal, M. (2014). Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. *Information & Management*, 51(5), 497–510.
- Pan, M.-J., & Jang, W.-Y. (2008). Determinants of the adoption of enterprise resource planning within the technology-organization-environment framework: Taiwan's communications industry. *The Journal of Computer Information Systems*, 48, 94–102.
- Premkumar, G., & Roberts, M. (1999). Adoption of new information technologies in rural small businesses. *Omega*, 27, 467–484.
- Ratten, V. (2012). Entrepreneurial and ethical adoption behaviour of cloud computing. *Journal of High Technology Management Research*, 23(2), 155–164.

- Silva, L. (2007). Post-positivist Review of Technology Acceptance Model ., 8(4), 255–266.
- Thong, J. Y. L. (2001). Resource constraints and information systems implementation in Singaporean small businesses. *Omega*, 29, 143–156.
- To, M. L., & Ngai, E. W. T. (2006). Predicting the organisational adoption of B2C e-commerce: an empirical study. *Industrial Management & Data Systems*, 106(8), 1133–1147.
- Tornatzky, L., & Klein, K. (1982). Innovation characteristics and innovation adoption-implementation: A meta-analysis of findings. *IEEE Transactions on Engineering Management*.
- Wu, W. W. (2011). Developing an explorative model for SaaS adoption. *Expert Systems with Applications*, 38(12), 15057–15064.
- Wu, Y., Cegielski, C. G., Hazen, B. T., & Hall, D. J. (2013). Cloud computing in support of supply chain information system infrastructure: Understanding when to go to the cloud. *Journal of Supply Chain Management*, 49(July), 25–41.
- Zhu, K., Dong, S., Xu, S. X., & Hally, M. (2006). Innovation diffusion in global contexts: determinants of post-adoption digital transformation of European companies. *European Journal of Information Systems*, 15(October), 601–616.
- Zhu, K., & Kraemer, K. L. (2005). Post-adoption variations in usage and value of e-business by organizations: Cross-country evidence from the retail industry. *Information Systems Research*, 16(1), 61–84.
- Zhu, K., Xu, S., & Detrick, J. (2003). Assessing drivers of e-business value: results of a cross-country study. *Proceeding of the 24th International Conference on Information System*, 1–13.
- Zhu, Y., Li, Y., Wang, W., & Chen, J. (2010). What leads to post-implementation success of ERP? An empirical study of the Chinese retail industry. *International Journal of Information Management*, 30, 265–276.

AGILE DRAFTING OF IT-OUTSOURCING CONTRACTS

A conceptualizing case study

Completed Research

Schlichter, Bjarne Rerup, Aarhus University, Aarhus, Denmark, brs@mgmt.au.dk

Storgaard, Kristian Kromann Reumert, Aarhus, Denmark, kst@kromannreumert.com

Abstract

The concept of ‘agility’ has become quite popular in the development of IT-artefacts and has created interest in the more general project management literature. The process of drafting complex contracts for large IT projects (service as well as software) is often done under time pressure and in several parallel tracks using different competencies. By the use of an illustrative case-study, this paper explores how scrum can be applied to enhance the process of drafting outsourcing contracts. The analysis indicates that the use of an agile method, such as Scrum, can be beneficial in this context with a minimal adjustment and that the elements of roles, processes and artefacts may lead to better coordination and efficiency as well as higher quality. The paper concludes with suggestions for further research and discussion of the findings.

Keywords: Contract drafting, Agility, Scrum

1 Introduction

Agile methods are highly used in the development of software and based on a belief that the project in a broad understanding continuously needs to be adjusted based on the learning acquired during the process. The drafting of contracts to frame the agreement between supplier and customer has shown to become more and more difficult, following the increased size and complexity of the IT-artefacts to be supplied, being software, hardware, services or a combination of these. One of the areas where specialists (lawyers) in charge of drafting these contracts have experienced problems is in the outsourcing of IS/IT-services (Lacity et al., 2010). In other contexts, the concept of agility has been seen as one of the means to handle complexity and changing requirements (Smith and Fingar, 2003), and thus this paper identifies issues in the current work practice and explores how the drafting of these complex contracts (the contractual framework and set of written agreements regulating the delivery of the services) can be conceptualised by using the artefacts from Scrum.

The offset of the paper is the work done by lawyers, who carry out or manage the actual drafting of complex IT and sourcing contracts. Structural and internal as well as external factors in the business of lawyers’ calls for cooperation and a more project based approach to providing legal advice and legal services (Storgaard, 2014b). It is seen that over the years, lawyers, especially in larger practices, have transformed from being generalists to become more specialized. The size and complexity of the cases naturally lead to the forming of groups of lawyers to meet the demand for resources in due time or the need for highly specialized competences.

The issues described above are prevalent during drafting of complex IT contracts where the actors involved face many of the same problems that the agile methods try to facilitate. Since agile methods have matured and are well documented in the development part of the procurement of software it

seems fruitful to try to adapt this concept to other parts of the procurement process, including the contract drafting process.

By assuming that the use of agile methods can improve both the process of drafting as well as the contract itself, the following Research Question has been formed:

How can drafting of complex IT-outsourcing contracts be conceptualized using the agile method Scrum?

This paper presents the contract drafting process in a practice of lawyers. First, the paper introduces the theoretical background on agile development, including Scrum, and presents how contemporary research in project management can frame these ideas. The research approach section describes the case, the data collection, and data analysis. The findings section presents our analysis of the actor's perceptions and use of agile elements in this context. Hereafter we discuss how the analysis addresses the research question and contributes to previous research. Finally, we summarize the conclusions of the paper and suggest how to proceed the research.

2 Theoretical Background

This section provides a theoretical and contextualised introduction to agile development and management of contracts.

2.1 Rethinking Project Management

The agile approach has been used in the general understanding on how to modernize project management, especially in the stream of research related to Rethinking Project Management (Svejvig and Andersen). Traditionally a project has been seen as a tool applied to a single assignment, focusing on meeting time and quality under the available resources. This single-track approach applies a logical and chronological way through a set of (more or less) well-defined tasks.

Newer ideas on management of projects introduce a plenteous of approaches, i.e. by understanding the project as a temporary organization, where learning, diversity, temporality, complexity, un-security and sociability are in play (Winter and Szczepanek, 2009). The increased complexity has led to the development of agile methods and to the acceptance of the need for a contemporary organization of projects that not only mechanically executes a process towards a narrow, product-oriented goal, but accepts projects as something business-like and value-creating (Davenport, 2013).

2.2 Agile development

Agile methods build on the assumption that the development of software can be complicated as well as un-predictable where upfront detailed planning is not always realistic. This class of methods shares the wish continuously to deliver value and learning (Cadle and Donald, 2008;p.78-80). Agile methods apply iterative and incremental development and hence divides the process into a number of smaller chunks – iterations – each ending with a delivery of functionality which is – in principle – ready for shipment. The detailed analysis of requirements and possibilities is done during the process at the beginning of each iteration (Schwaber and Beedle, 2002). Thus, no detailed requirement specification exists initially. The final delivery is decided during the iterations by exploiting the experience and knowledge achieved. This approach assures quality as well the possibility to adjust the project in a volatile environment thus reducing risk (Cervone, 2011, p.18-22). The principles constituting the agile concept are given in the 'agile manifesto' (Beck et al., 2001) and establish, among 11 others, that the highest priority is to provide the customer with early and continuous delivery of valuable software, illustrating that agility is not about abolishing traditional practices, but moving the focus from a 'locked' requirement specification to the business goals of the customer (Martin, 2003).

2.3 Scrum

Scrum is an iterative and incremental development approach, where planning is concurrent to the development activities and the work is divided into smaller chunks called sprints. Each sprint is planned to be self-contained, leading to a new running version on the road to the final software product (Jakobsen and Sutherland, 2009). Scrum is a framework to be used when developing complex products and based on the theory of empirical process control theory (empiricism), assuming that knowledge comes from experience and that decisions must be based on what is known (Sutherland and Schwaber, 2013) to optimize predictability and reduce risks.

Empirical process control is based on three pillars (Schwaber and Beedle, 2002):

Transparency (The participants should have insight and be able to understand the process.)

Inspection (This must be controlled on a recurrent basis)

Adaption (if needed, adjustments to meet the goal must be done)

Scrum is based on three main sets of elements: Roles, Event/Ceremonies and Artefacts. The elements have specific goals and are, as a whole, essential to the use and success of Scrum (Schwaber and Beedle, 2002). To establish a foundation for conceptualization the main set of Scrum Artefacts are presented below.

2.3.1 Roles (The Scrum Team)

The Scrum team is self-organizing and interdisciplinary, thus having the competences needed to act independently (Sutherland and Schwaber, 2013). The team consists of a Product Owner, a Development Team and a Scrum Master and is designed to optimize flexibility, creativity and productivity.

The **Development Team** is a group of software developers assigned to finalize the Sprints by delivering a potential shippable chunk of functionality (SW) respecting the budget.

The **Product Owner** is overall responsible for (maximizing the value of) the product and maintaining contact to the 'customer' and other stakeholders, thus being able to answer all questions.

A **Scrum Master** is a "gatekeeper" ensuring that the Scrum Team adheres to Scrum theory, practices and rules, as a of "serving leader" for the team. The Scrum Master offers coaching of the team and protects against disturbance from outside.

2.3.2 Ceremonies

Scrum is based on a set of fixed meetings and rituals to ensure regularity, and minimize obsolete meeting activity (Rubin, 2012,p. 7):

The **Sprint** is a delimited period of time (max one month) where the development of a chunk of software, in principle shippable, takes place. The idea is that no changes in neither the goals of the Sprint, the quality standards nor in the composition of the team are allowed.

The **Daily Scrum** is a <15 minute meeting used to synchronise activities and compose a plan for the next 24 hours. Achievements since last Scrum are inspected and each member has to account for:

“What did I do yesterday that helped the Development Team meet the Sprint Goal?”

“What will I do today to help the Development Team meet the Sprint Goal?”

“Do I see any impediment that prevents me or the Development Team from meeting the Sprint Goal?” (Sutherland and Schwaber, 2013, p.10)

Based on this, the Development Team can evaluate the progress towards the Sprint Goal and at the same time evaluate the progress in relation to the Sprint Backlog (e.g. by using the Burn Down Chart).

The **Sprint Review** is done by the Scrum team at the end of each Sprint with the purpose to inspect the software produced and update the Product Backlog. Often other stakeholders, such as management and sales, other developers and support staff take part to get insight in status and to give feedback. The Sprint Review is, as other meetings meant to be limited in time (e.g. four hours at a one-month-sprint).

The **Retrospective** is done by the end of each Sprint and is a review of achievement in the Sprint and updating the Product Backlog. The team and other relevant actors (management, sales staff, other developers, support functions etc.) take part in the short meeting to create a plan for improvements as well as coordination.

2.3.3 Documents

In addition to Roles and Ceremonies, Scrum applies Documents (sometimes called artefacts), these are tools designed to assure visibility and transparency for production in the Scrum Team.

The **Product Backlog** is a sorted and prioritized list of desired functionality in the final software product and hence the central, source of information about what should be produced and when. The product owner is responsible for the Backlog, including content, availability and prioritizing. The backlog is a living document based on initial insight and updated according to new requirements, experiences etc. (Rubin, 2012p. 12ff). All items contain a description, a priority and an estimate. The higher the priority, the more details. User stories ('as user of XXX, I would like YYY, so that I can ZZZ') are often used to describe functionality in the Product Backlog.

The **Sprint Backlog** consists of the Product Backlog items assigned to the actual Sprint, and is as such a plan how to deliver the product incrementally and how the work should be done by the Development Team to transform Product Backlog Items into (parts of) a product.

The **Definition of Done** is a set of parameters to establish a shared understanding of when a delivery, a function or a product can be seen as 'complete' and hence to be released. In parallel the Burn Down Chart graphically indicates how well the team follows the plan. It is considered important that the actors have a shared understanding when something is done.

2.4 Scrum in other settings

Initially Scrum was developed to be used for software development, but recently several examples of use in other domains has emerged (Cadle and Donald, 2008), i.e. in universities, law defences, the car industry and development of mechanical products (REYNISDÓTTIR, 2013). At Roskilde University in Denmark, Scrum was used to rapidly produce a policy document. The evaluation indicated that Scrum could be used for development of reports and that the participants were provided with an overview of the project not seen before and should not wait for individual deliverables (Pries-Heje, 2012, p.130).

Apparently no examples on the direct use of Scrum on projects related to legal advice of drafting of contracts can be found (Goel, 2013), even though the idea of applying Scrum or agile methods within the legal industry is not totally novel, examples has been seen in litigating (Terrett, 2010). Drafting of contracts as such is well discussed (Pumphrey, 2009).

3 Current Practices

This section presents the case and its related context followed by an explanation of how we collected and analysed data. The case study approach is in terms with Cavaye's single case with interpretive use of qualitative data for discovery (Cavaye, 1996). This interpretive research approach allowed us to

investigate how the concepts of Scrum can be applied to contract drafting for complex IT-projects in their organizational and cross-cultural context and thus open to several interpretations by organizational actors, but also to us as researchers (Klein and Myers, 1999, Walsham, 1995, Walsham, 2006). One of the authors is an employee of the law firm used for the case study.

3.1 The case

The law firm, on which this case study is build, is one of the largest Danish law firms. A major part of the legal services are focused on transactions such as procurement of companies or services and hence drafting of complex contracts.

The lawyers' work is often centred on a case which can be a dispute (i.e. in form of a court case), a transaction (procurement, i.e. an IT project or something related to outsourcing) or more general advisory (i.e. on the structure of a company).

A new case will typically be staffed with a group of specialists (often from different departments. The case will always be anchored to a responsible partner (owner) and often a senior associate may be appointed as project manager to take care of progress, coordination etc. In some cases additional project support staff may be assigned to the case.

In case of complex IT and outsourcing contracts, there is a need for specialists in e.g. outsourcing, IT and terms of employment. They each develop their part of the contract in a process that could take weeks or even months. During this process, a large amount of information is collected and developed. A complex IT or outsourcing contract typically consists of a main contract and a large number of annexes. The main contract is the basic legal framework, where the annexes normally contain more detailed regulation or descriptions of different parts of the delivery, e.g. a catalogue of services, service level agreements or regulations of transfer of employees. Since all contributions will eventually be a part of same contract, there is a need to align them regarding e.g. terminology as well as assuring that an issue is not regulated (differently) more than one place.

Drafting of this type of contracts hence faces a set of challenges:

- a) The work takes place in a project group with different employees not coming from the same department, not necessarily having worked as a team before (or even having limited experience from similar cases).
- b) The initial level of knowledge will often be low and increasing during the project. Knowledge comes from many sources and not always in a structured form. Knowledge and experience is generated continuously, and there is an ongoing need of adaption, coordination and prioritizing of the different deliveries.
- c) Requirements to the contract may change often and with short notice. Sometimes parts of the contract must suddenly be drafted differently or certain aspects regulated otherwise, e.g. due to commercial demands.
- d) Time pressure can often be severe

If not taken care of, the above challenges may have negative impact on the writing speed as well as on the quality of the resulting contract. Ultimately this may impose a risk on the client and similarly liability for the lawyers.

3.2 Data Collection

The data collection included document studies and individual semi-structured interviews with team members and the management of the law firm on which the case study is based. The majority of these interviews, as well as insider observations, are documented in two reports compiled by one of the au-

thors (Storgaard, 2014b, Storgaard, 2014a). In addition to forming the base for the analysis in the present paper, the interviews are seen as pilot interviews for the next phase of the research, since they furthermore provided an understanding of the environment and the challenges faced by the organization and helped identify additional candidates for interviewing.

3.3 Data Analysis

To understand how an agile project model can be applied for drafting of complex contracts, the analysis will be centred around a conceptualization of Scrum on the contracting process. Scrum is initially constructed with software development in mind and hence uses terminology and frames from the IT domain. To apply Scrum and evaluate the fit in a contract drafting context, an ‘interpretation’ is needed.

The analysis is initially done by applying the Scrum artefacts on the actors, tools and processes used during contract drafting. For each of the artefacts it is then evaluated whether it can be applied as it is or has to be adapted. In the latter situation it is then decided if the adaption should be in the Scrum framework for the contract drafting process.

4 The conceptualised framework

In this section, we present how our observations and insight from the case study can be conceptualized using the Scrum framework. The analysis has the contract drafting process in focus.

4.1 Problems related to the existing drafting practice

The interviews revealed a set of challenges of the present drafting practice of large and complex contracts (Storgaard, 2014b):

Exchange of information: A centralized flow and collection of the continuous influx of new information are difficult. As a result there is a major risk that lawyers end up working in siloes.

Task Management: The many sub-tasks are often changed, and tasks / delegation of responsibility are done structurally by using e-mails or oral agreements.

Lack of project management capabilities: The lawyers are not educated in the practice of projects and are thus lacking potential benefits and synergies.

Diversity of tools and processes: The work practice can be quite fragmented and may be based on many different tools and processes, even when applied on similar tasks.

Process Management and communication with clients: A single and straightforward system where the client can stay updated on progress is lacking.

Accounting and Estimation: Within the legal industry it is common to invoice based hours spent, which do not fully support the working in projects or even co-operation.

The analysis below indicates how some of these problems can be alleviated using a Scrum approach.

4.2 Roles

In relation to an outsourcing contract, the senior associate (senior-lawyer) who is project manager will take the role as **Product Owner**, this is in line with the present situation where this role is to assure the value of the work and coordinate ‘upwards’ with the stakeholders. A new task will be the responsibility for administration of a centralised Product Backlog and hence defining the acceptance criteria for the individual backlog items. This requires more time to cope with the annexes than today, but will also assure a qualified overview – which today is often not fully present. Initially we refuse sugges-

tions of (partly) appointing the Partner as product owner, since a partner is also seen as part of the management.

The Development Team is the group of lawyers, making the actual drafting of the outsourcing contract. Similar to SW development, this team should be interdisciplinary in nature (legal specialties and branches). Some adjustments are needed though, as the senior-associate in the present setup are both part of the production team and have managerial responsibilities for the team, and this is not in line with the Scrum principles. This can be partly solved by letting the Scrum team take over the development of the framework agreement (which is often responsibility of the project leader).

The present setup is without **Scrum Master**. Since it is not required that a Scrum Master has a deep professional insight (in other words: being able to take part in the production) (Rubin, 2012, 1.4285), and it is recommended that the Scrum Master is not a part of the production team, it would be feasible to let one of the non-lawyers from the project support staff take this role. This is in consistence with the present situation where the project assistants support the project, and it would also let the assistants support the Product Owner by identifying procedures for efficient management of the Product Backlog.

4.3 Ceremonies

The ceremonies are important to secure transparency and provide possibility for inspection and hence adjustment of processes and deliveries. In an outsourcing contract drafting context this can be contextualized as explained below.

The Scrum Sprint is a novelty even though sub-deliveries are not unknown. Often the contract is delivered including annexes in two or three packages according to importance. To follow the idea of Sprint, the packages must be divided into even smaller chunks and timed after the actual project. Since the duration of the project is limited to a few months, a sprint lasting one to two weeks seems feasible.

The Sprint is initiated with a Sprint Planning Meeting, where elements from the product backlog are prioritized into the sprint backlog explicitly as to what the Sprint will deliver and how.

This type of planning is not seen in the present context, mainly caused by the introduction of the shorter sprints. Presently this division of tasks between participants (often in the form of assigning the drafting of different annexes) is done in the start-up of the process. The analysis indicates the concept of the sprint, and sprint planning can facilitate realistic and aligned goals for each sub-delivery. As a (positive) side effect, the client will experience more deliveries and ongoing progress in the project.

Daily Scrums will be a relevant activity improving transparency among the members of the production (development) team. All members will know the status of the project and problems occurring in the other tracks. This will reduce the identified risk of working in silos and enhance the efficiency (as is also seen in the Roskilde University case (Pries-Heje, 2012,p. 130)) because the team members should not wait for each other. In the light of the shorter sprint, it could maybe be beneficial to introduce more than one Daily Scrum.

Sprint Review is one of the activities that must be adjusted to be feasible in an outsourcing context. A sprint lasting 1-2 weeks will easily result in +100 pages of contract and annexes. Due to the nature of the product, it is not easy to demonstrate this and receive feedback during a two hour meeting (contrary to software, where the new functionality could often rather easily be demonstrated). Therefore – the sprint review must be adapted to the context. Either: i) the participants must read the documents beforehand, letting the meeting be limited to discussion or feedback or ii) another type of presentation. Suggestion i) is not realistic due to the workload without a comparable anticipated benefit. The practical solution may be that the development team presents highlights and/or items, where feedback is particularly needed, followed by further in-depth reading. It was also suggested to invite the responsible partner and other relevant specialists not involved on a daily basis.

The Sprint Retrospective is used to evaluate the work done in the Sprint by reflecting on persons, relations, processes and tools summarized in a plan to improve the contextualized Scrum Process. This ceremony is also quite relevant for the outsourcing contract drafting process as well, but may be met with scepticism since process evaluation is not a part of the ‘culture’ in this environment. In addition, the time pressure will somehow prioritize production in favour of reflection, given the often narrow deadlines.

4.4 Scrum documents

The Scrum documents are artefacts documenting status, progress or work to be done.

The Product Backlog will be a central tool in the contract drafting process. In the present process a simple list of documents with their status (level of completeness) are normal. By introducing the Scrum approach, additional dimensions such as prioritizing and estimating are added. This will for sure improve the visibility and establish a better basis for planning. The use of User Stories could also be applied on contract drafting. Just as software ultimately is to be used by the users, then an outsourcing contract is to be used by the client in respect of the client’s customer or service provider. User Stories can describe ‘functional’ requirements for regulations seen in a client perspective. Perhaps it would be even easier for the client to describe the needs, and this would force the lawyer to stick to the business goals of the client. Some lawyers are worried by the higher workload imposed when developing and maintaining the Product Backlog.

The **Spring Backlog** and its progress, as documented in the **Burn down Charts**, can also be beneficial during the contract drafting process. In case the Product Backlog contains wishes about a specific regulation (e.g. a guaranty or specific service), the breakdown of work into items in the Sprint Backlog tells explicitly where in the contract complex this is to be done and indicates the interdependencies to be elaborated on. Since deadlines are quite firm in this context, it is very important to get an ‘early warning’ in case of potential problems to meet these or to take other measures to align expectation with the client.

Definition of Done, also relevant for the release backlog, would be a novel thing in contract drafting. Even though a shared understanding (maybe implicit) about the quality or type a contractual document shall have, no checklists exist. A Definition of Done could be used for quality assurance and could be adjusted to the different contracting types. A structured approach to quality assurance *during* the contracting process, instead of after, is expected to reduce the number of errors in the finalized contracts.

| # | Scrum | Contracting | Observation |
|---|-------------------|---|---------------------------------------|
| | <i>Roles</i> | | |
| 1 | Product Owner | The senior lawyer already acting as project manager. | |
| 2 | Scrum Master | Assistant from the Project Center | A non-lawyer |
| 3 | Development Team | A group of Lawyers from the production | Interdisciplinary |
| | <i>Ceremonies</i> | | |
| 4 | Sprint | Smaller chunks than today’s ‘packages’ | A novelty |
| 5 | Daily Scrum | Quite relevant, introduce transparency etc. | |
| 6 | Sprint Review | Changed to e.g. presentation and discussion of pre-read documents | Difficult to present a ‘running’ item |

| | | | |
|----|--------------------|---|---------------------------|
| 7 | Retrospective | To be introduced, relevant | Cultural barrier. |
| | <i>Documents</i> | | |
| 8 | Product Backlog | A simple list of documents and their status to which are added: prioritization and estimating | Addition of user stories? |
| 9 | Release Backlog | Not conceptualized. | |
| 10 | Burn Down Charts | To be introduced | novelty |
| 11 | Definition of Done | To be introduced. It would be beneficial to introduced a set of shared standards | novelty |

Table 1: Conceptualization

5 Discussion

The case description accounts for problems related to the present contract drafting process. Some of these problems resemble experiences in the development of software, where agile methods seem to alleviate these problems. A few examples of the first tentative attempts to apply Scrum in other domains have been identified. The analysis shows how the contract drafting process (of complex outsourcing contracts) can be conceptualised using Scrum artefacts. For each artefact, pro et cons are identified and discussed. Scrum is a framework to be used in its full extent to give the best benefit (Sutherland and Schwaber, 2013, p.16) which questions if the use of Scrum on drafting of complex IT- or outsourcing contracts fully alleviates the challenges in the contract drafting process. The answer seems to be confirmative:

- 1) The **Scrum Roles** can be used and the distinct division of roles and responsibilities seems beneficial. It is expected that as Product Owner, the project responsible lawyer will work more with project management (in opposition to production) – but this extra workload will be balanced with a better overview and higher efficiency in the Development Team.
- 2) The **Scrum Rituals** can be used with some adaption to the contract drafting process. It seems that the use of Scrum (sprint, scrums etc.) will lead to better coordination, improved exchange of information as well as ongoing learning. This is consistent with experiences from other domains (Pries-Heje, 2012, pp129-135).
- 3) The **Scrum Documents** can also be applied to the contract drafting process. The use of Product Backlog and Definition of Done (and maybe even a Burn Down Chart) is expected to improve quality, coordination and conformity. The use of Scrum during contract drafting seems to alleviate some of the problems, especially regarding sharing information, task management, processes and tools and somehow also accounting.

6 Conclusion and further research

This paper answers the research question: *How can drafting of complex IT-outsourcing contracts be conceptualized using the agile method Scrum?* The contract drafting process has been successfully conceptualised using Scrum concepts as shown in table 1. Suggestions on how to change the drafting process in a Scrum conceptualisation are given.

During the following research project we will, with offset in the contributions from the present paper, broaden our literature study to identify contemporary research that can support the development of our

findings. It could also be beneficial to re-visit our case and collect more data to challenge and discuss our findings, especially focusing on different actors (including the client side).

References

- BECK, K., BEEDLE, M., VAN BENNEKUM, A., COCKBURN, A., CUNNINGHAM, W., FOWLER, M., GRENNING, J., HIGHSMITH, J., HUNT, A. & JEFFRIES, R. 2001. Manifesto for agile software development.
- CADLE, J. & DONALD, Y. 2008. *Project Management for Information Systems*, Essex, England, Pearson Education Limited.
- CAVAYE, A. L. M. 1996. Case study research: a multi-faceted research approach for IS. *Information systems journal*, 6, 227-242.
- CERVONE, H. F. 2011. Understanding agile project management methods using Scrum. *OCLC Systems & Services: International digital library perspectives*, 27, 18-22.
- DAVENPORT, T. H. 2013. *Process innovation: reengineering work through information technology*, Harvard Business Press.
- GOEL, S. 2013. Agility - Looking Beyond Software Development. *Scrum Alliance Community*.
- JAKOBSEN, C. R. & SUTHERLAND, J. Scrum and CMMI Going from Good to Great. Agile Conference, 2009. AGILE '09., 24-28 Aug. 2009 2009. 333-337.
- KLEIN, H. K. & MYERS, M. D. 1999. A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly*, 23, 67-93.
- LACITY, M. C., KHAN, S., YAN, A. & WILLCOCKS, L. P. 2010. A review of the IT outsourcing empirical literature and future research directions. *Journal of Information Technology*, 25, 395-433.
- MARTIN, R. C. 2003. *Agile software development: principles, patterns, and practices*, Prentice Hall PTR.
- PRIES-HEJE, J. 2012. *Project Management Multiplicity*, Samfundslitteratur.
- PUMPHREY, J. H. 2009. Thinking like a Writer: A Lawyer's Guide to Effective Writing and Editing by Stephen V. Armstrong and Timothy P. Terrell. HeinOnline.
- REYNISDÓTTIR, Þ. 2013. *Scrum in Mechanical Product Development. Case Study of a Mechanical Product Development Team using*. Master, CHALMERS UNIVERSITY OF TECHNOLOGY.
- RUBIN, K. S. 2012. *Essential Scrum: A practical guide to the most popular Agile process*, Addison-Wesley.
- SCHWABER, K. & BEEDLE, M. 2002. *Agile software development with Scrum*, Prentice Hall Upper Saddle River.
- SMITH, H. & FINGAR, P. 2003. *Business Process Management: The Third Wave*, Tampa, USA, Meghan-Kiffer Press.
- STORGAARD, K. 2014a. Agil Metode Anvendt på IT-Kontrakter. Aarhus University.
- STORGAARD, K. 2014b. IT-Mæssig undersøttelse af samarbejde og projektarbejde i en advokatvirksomhed. Eksamensopgave i faget "IT strategi i Praksis": Aarhus Universitet.
- SUTHERLAND, J. & SCHWABER, K. 2013. The Scrum Guide.

- SVEJVIG, P. & ANDERSEN, P. Rethinking project management: A structured literature review with a critical look at the brave new world. *International Journal of Project Management*.
- TERRETT, A. 2010. Do Litigators "scrum" ? *Slaw*.
- WALSHAM, G. 1995. The Emergence of Interpretivism in IS Research. *Information Systems Research*, 6, 376-395.
- WALSHAM, G. 2006. Doing interpretive research. *European Journal of Information Systems*, 15, 320-330.
- WINTER, M. & SZCZEPANEK, T. 2009. *Images of projects*, Gower Publishing, Ltd.

MOBILITY AND SECURITY IN THE NEW WAY OF WORKING: EMPLOYEE SATISFACTION IN A CHOOSE YOUR OWN DEVICE (CYOD) ENVIRONMENT

Complete Research

Arjan de Kok, Department of Information and Computing Sciences, Utrecht University,
Utrecht, The Netherlands, a.dekok@uu.nl

Yvette S. Lubbers, Department of Information and Computing Sciences, Utrecht University,
Utrecht, The Netherlands, i.lubbers@uu.nl

Remko W. Helms, Department of Information and Computing Sciences, Utrecht University,
Utrecht, The Netherlands, r.w.helms@uu.nl; Faculty of Management, Science &
Technology, Open University, Heerlen, The Netherlands, remko.helms@ou.nl

Abstract

The consumerization of IT, known as Bring Your Own Device (BYOD), is an inevitable component in the future IT infrastructure of organizations. It is not the question if employees will use consumer IT products for their work, but how and under which conditions. The use of personalized mobile devices may be beneficial for both the employee and organization, but the concern of IT executives, on corporate data residing on uncontrolled mobile devices, is often leading to a restrictive policy. Giving employees the ability to choose from a variety of secure devices, at the expense of the organization, Choose Your Own Device (CYOD), may well bring the best of two worlds. In this research 126 employees at four multinational organizations were surveyed on their perception of usability and satisfaction of devices for their knowledge tasks. The outcomes were matched against a Risk Assessment on seven identified IT threats. The results show that a majority (52%) believes their performance would improve, when given the ability to choose a device of their own. The Risk Assessment shows that IT security risks do not need to increase, provided that the proper security policies are in place. This implies that the performance and satisfaction of employee can improve in a secure CYOD environment.

Keywords: New Way of Working (NWOW), Choose Your Own Device (CYOD), Bring Your Own Device (BYOD), Consumerization.

1 Introduction

In the new world of work the use of consumer IT for business purposes, consumerization or Bring Your Own Device (BYOD), has seen a tremendous flight in the past years (Gillett, 2012; Citrix, 2013). Employees perceive personal devices to be more useful, more powerful, easier to use, and more fun than enterprise IT, and often they are (Harris et al., 2012). Personal devices have become inexpensive and the software apps are low cost or for free. On the other side, IT executives have concerns, mainly about data security, when employees view and use corporate information on their own mobiles, tablets and other personal devices. Also, BYOD confronts IT departments with a wide variety of software platforms that are used to connect to the corporate network, on devices that are renewing at a much faster pace than upgrades that were rolled out in the past. The reaction is often a push towards tight control, imposing restrictive, and often performance-taking, software on employees' devices. The question that is now raised by employees is: "Should I be the one to pay for working more effective and pleasurable, while receiving corporate control over my privately owned hardware?" This results in a situation that makes both parties feel uncomfortable.

A solution that seeks to find a ‘middle-way’ in this impasse is Choose Your Own Device (CYOD). Choose Your Own Device enables employees to choose, against no personal costs, the devices that they feel suit them best in the tasks they need to perform, whilst allowing the organization to supply enterprise-controlled technology. Having the benefits of both worlds, CYOD is growing in popularity, especially in larger organizations. Where there is existing research on BYOD, research in the field of CYOD policies, especially in the light of IT security, and in the context of the New Way of Working, is scarce if not at all absent.

The research question is: Can a CYOD policy contribute to a perceived improvement in employee performance and satisfaction, in a secure way? In this research 126 employees were surveyed at four large organizations, that had chosen for a CYOD policy, whilst seeking the optimum of IT security and user satisfaction. The context of this CYOD environment (at least for the Dutch divisions of these companies), was the New Way of Working. The following chapter (2) briefly describes the context of the New Way of Working and CYOD, the tasks of knowledge workers and threats in IT security. The research method is explained in chapter 3 as well as the Technology Acceptance Model that is used for the determination of the device usefulness and user satisfaction. Chapter 4 discusses the research results. This leads to a number of conclusions and recommendations for future research in chapter 5.

2 Theoretical background

2.1 The New Way of Working

Where in the past many authors e.g. Hammer & Champy (1993) envisioned a ‘New World of Work’, with information technologies as rule-breaking for the way business processes would change, the last decade has shown an increase in pace in which new ways of working are being adopted in organizations. Bødker & Christiansen (2002) were one of the first to observe that ‘new work is characterized by a mobile, networked technology, project-managed organization, and new office designs. The office designs are explicitly motivated by the wish to facilitate creativity, knowledge sharing and communication, carried out across a variety of settings: office, home, airports, coffee shops and cars’ The creation of new office spaces that are breaking with all traditional rules and design concepts is probably one of the most visible effects of the New Way of Working (NWOW). Offices transform from dull production facilities to inspiring meeting places, in which no effort is spared to create a new sense and experience of work (Waber et al., 2014). At the same time employees enter into new working relations in which they have the freedom to decide when and where to work, and become responsible for their results instead of being measured by their ‘presenteeism’ at the office (Johns & Gratton, 2013).

Baane et al. (2010) add: ‘The work principles of The New Way of Working give maximal freedom to employees, on the basis of mutual trust. This trust is expressed in the freedom that employees have for carrying out their work in ways, times and locations that suit them best. The employees are evaluated based on their personal or team contribution, rather than their presence. Thus the employees can engage in a working relationship that fits in terms of ambition, skills, lifestyle or stage of life’. The context of NWOW can be divided into three dimensions: Bricks, Bytes and Behavior. (1) Bricks, the physical dimension, addresses all aspects of the physical work environment, (2) Bytes, the technological dimension, that addresses all aspects concerning the use and application of ICT, and (3) Behavior, the personal dimension, which addresses all aspects concerning the manager-employee relationship and the way the employee works and experiences his or her work.

2.2 Knowledge tasks

The work principles of NWOW are best applied in the work environment of the ‘knowledge worker’ (Greene & Myerson, 2011). The term knowledge worker is not new: already in 1969 Drucker used the term knowledge worker for ‘the man or woman who applies productive work ideas, concepts and information rather than manual skill or brawn’. The question is: which tasks are performed in the work environment of the knowledge worker, and which device would suit the execution of this task well, in

the perception of the knowledge worker? Reinhardt et al. (2011) researched the roles and actions knowledge workers perform. In their literature review they analyzed all the knowledge actions described by different authors (e.g. Davenport & Prusak, 1998) and combined them to one coherent list of knowledge actions. These tasks were used in the Employee Survey in this research. For an overview of the knowledge tasks and their description see Appendix 1.

2.3 Consumerization of IT

Mobility is an important aspect in the vision of the New Way of Working to work anywhere and anytime. For employees it is important to work with the devices that are best suited for their work, adding the 'work with anything' aspect to working anywhere and anytime. Moschella et al. (2004) were probably the first ones to coin the term Consumerization of IT (Ruch & Gregory, 2014). They concluded employees were often so frustrated with the existing IT infrastructure, that they chose to bring and use their own devices for their work. The work with personal consumer devices for business means is since called IT consumerization or Bring Your Own Device (BYOD). Giddens & Tripp (2014) define BYOD as 'the use of personal devices at work, on the workplace, to complete work-related activities'. Ingalsbe et al. (2011), Holtsnider et al. (2012), and Harris et al. (2011) use similar definitions for the dual use of devices for private and business purposes. The use of consumer IT devices for business purposes is expected to contribute to work performance and greater autonomy for employees (Niehaves et al., 2012, 2013). Murdoch et al. (2010) and Harris et al. (2011) add that employees using the technology of their own find it easier to use and important for their job satisfaction.

Though many companies struggle with this phenomenon, and often do not have a BYOD program in place, the reality is that employees already bring their personal devices to work (Gillett, 2012; Citrix, 2013). Forrester Research found that 52% of the information workers use three or more devices for work (Gillett, 2012). They predict that by 2016 there will be 760 million tablets in use, most for use both at work and at home (Gillett, 2012). As companies reap the benefits, but employees pay the cost, of the improved work performance, a number of companies decided to sponsor the use of personal devices. Sometimes this sponsoring goes under the condition of allowing company security controls on one's personal device. In particular the security aspects of protecting business data fragmentation on a broad range of personal devices is challenging to implement. ICT managers however realize this trend cannot be stopped, and therefore needs to be managed. Because of the security aspects, a number of organizations consider a Choose Your Own Device (CYOD) policy in which employees are allowed to choose from a range of mobile devices with pre-installed security management software in place, at no personal cost.

A CYOD policy can optionally be combined with a BYOD policy, for instance when users agree to have security software installed on their personal device as well, but often it is restrictive in the form of a Don't Bring Your Own Device (DBYOD) policy. In this case personal devices are not allowed to connect to the corporate network. In practice this means that employees in a DBYOD environment can only access the restricted guest network from their own device.

2.4 IT threats

An IT risk can be defined as the damage or impact an event or threat will cause, against the chance or probability of its occurrence (Baskerville 1993; Peltier, 2005). The chance of occurrence may be both erroneous human actions and attackers who attempt to abuse weaknesses in technical solutions. Mobile devices e.g. notebooks, tablets and smartphones are often used outside the corporate network. Mostly users are able to install software or apps, and connect to multiple public domains. Often users do not realize the potential damage this may cause. Morrow (2012) found that around 40% of the employees admit they do not update their (security)software, while unauthorized access to and information theft from endpoints has increased by malware, key loggers and cyber-attacks. Even when anti-virus software is present, mobile malware can be effective, and steal user credentials.

Security risks constantly change over time, making research in this area time-bound. Whitman (2003) identified twelve categories of IT security threats of both human and technical ground. In the light of this research some categories were identified as not applicable (e.g. force of nature), or not essentially different for the types of researched devices. The results was the following list of seven IT security threats that were identified for this research:

| | Threat | Examples |
|---|--|--|
| 1 | Act of Human Error or Failure | accidents, employee mistakes |
| 2 | Compromises to Intellectual Property | piracy, copyright infringement |
| 3 | Deliberate Acts of Espionage or Trespass | unauthorized access and/or data collection |
| 4 | Deliberate Acts of Theft | illegal confiscation of equipment or information |
| 5 | Deliberate Software Attacks | viruses, worms, macros, denial of service |
| 6 | Technical Hardware Failures or Errors | equipment failure |
| 7 | Technical Software Failures or Errors | bugs, code problems, unknown loopholes |

Table 1. *IT security threats (Whitman, 2003)*

There are roughly three mechanisms to cope with IT security risks: (1) authentication, (2) network security and (3) device security. (1) Authentication is ‘the process of determining whether someone or something is, in fact, who or what it is declared to be’ (Rouse, 2007). When a user is authenticated, identity and access management can be applied. This security discipline ‘enables the right individuals to access the right resources at the right times for the right reasons.’ (Gartner, 2015). (2) Network security is the policy to prevent unauthorized access to the corporate network. Almost all corporate laptops nowadays use a VPN connection to access corporate data from an external connection. Information from a virtual private network is securely transported over a public network by encrypting the data to keep it confidential (Govcert, 2009). (3) Device security. This can be enforced using software such as a Mobile Device Management (MDM) tool. This software is installed on the mobile device and encrypts the (corporate) data. It enables the employer to monitor the entire device, push software updates, and remotely kill data stored on the device in case of loss or theft (Gajar et al., 2013). Ideally, organizations are able securely deliver corporate data to employees, without interfering with their access to personal apps and data. However, the ability to separate corporate data from personal data on a mobile device has its limits. E.g.: Was the picture taken by the camera a business whiteboard or holiday picture? Information security will therefore always be a balancing act of business interest versus personal freedom.

3 Research method

3.1 User acceptance models for IT

To determine the user acceptance of information technology, multiple models have been developed. In this section two models are discussed: the Person-Artifact-Task (PAT) model from the Flow theory, and the Technology Acceptance Model (TAM).

The Flow theory originates from Psychology. The psychologist Csikszentmihalyi (1975, 1988, 1990) found that people can be so absorbed in an activity, such as chess playing or rock climbing, that they excel in performance and lose track of time, without being aware of it. When personal computers were introduced, the Flow theory was used to address user experiences in computer-mediated environments (CMEs), such as the satisfaction and acceptance of information technology (Ghani, 1991). Based on the Flow theory, Finneran & Zhang (2002) defined the Person-Artifact-Task (PAT) model, in which activities are broken down into tasks and artifacts (tools), that need to be mastered by the user. The

likelihood of an optimal (flow) experience depends on the interplay between the person, the task and the artifact. Kiili (2004) presents a framework of the factors in each stage of flow with the components of the PAT model, see Figure 1.

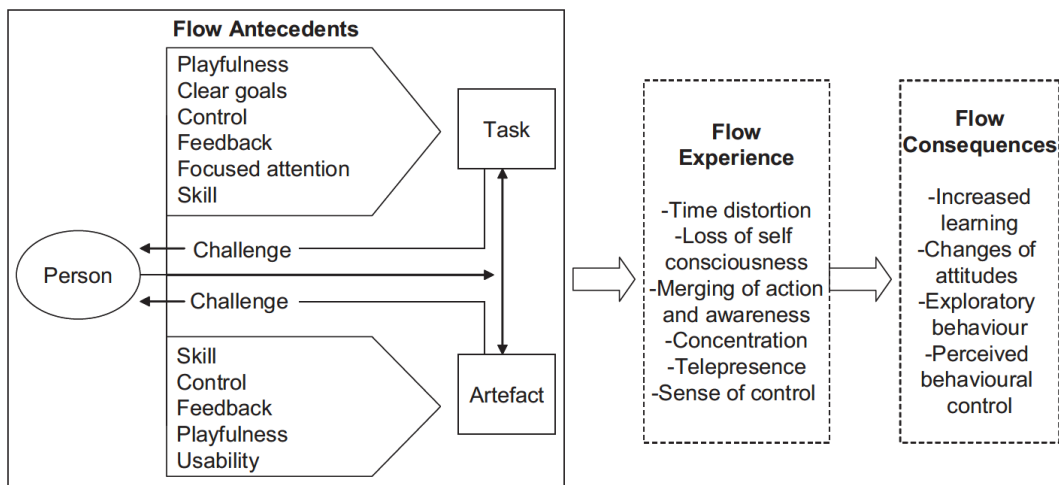


Figure 1. Person-Artifact-Task (PAT) model (Finneran and Zhang, 2002, Kiili, 2005)

In this framework the antecedents Speed and Ease of use (Skadberg & Kimmel, 2004), are combined as the Usability factor. Perceived ease of use (PEOU) is an established and validated construct in MIS literature (Davis, 1998; Venkatesh & Davis, 1996, 2000).

Based on the Theory of Reasoned Action (TRA) of Ajzen & Fishbein (1980), which suggests that people form intentions to adopt a behavior or technology based on their beliefs about the consequences of adoption, Davis (1998) builds the Technology Acceptance Model (TAM). In this model two major variables determine an individuals' information system acceptance; Perceived usefulness and Perceived ease of use. In the extended Technology Acceptance Model (TAM2), Venkatesh & Davis (2000) incorporate several additional attributes that influence system acceptance, e.g. Output quality. Figure 2 shows the extended Technology Acceptance Model.

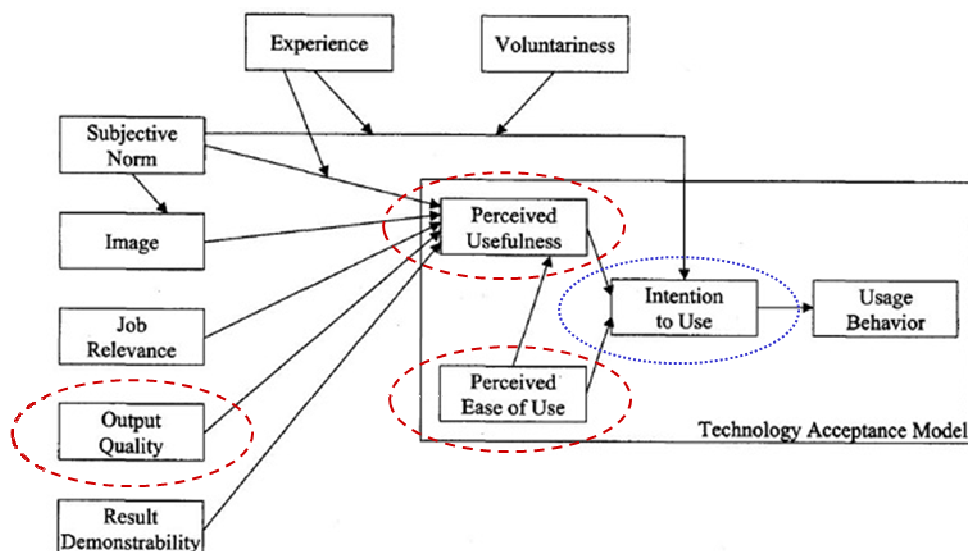


Figure 2. Extended Technology Acceptance Model (TAM2) (Venkatesh and Davis, 2000)

In Figure 2 the first three constructs, that are used in this research, have been circled with a red dashed line. They are defined by Venkatesh & Davis as follows: Perceived usefulness is the extent to which a

person believes that using the system will support or enhance his or her work job performance. Perceived ease of use is the extent to which a person believes that using the system is or will be free of effort. Perceived usefulness is influenced by Perceived ease of use because, other things being equal, the easier the system is to use the more useful it can be. Output quality is the degree to which a person believes the system performs his or her job tasks well.

The fourth construct that is used in this research is (perceived) satisfaction. This construct is not as such in the TAM2 model, but it is related to the Intention to use, which therefore has been circled with a dotted blue line. Wixom & Todd (2005), who tried to combine the attributes from user satisfaction literature with the Technology Acceptance literature, warn that user satisfaction is limited in its ability to predict system usage. The question is therefore what leads to satisfaction and intended system use.

Giddens & Tripp (2014) suggest that device self-efficacy, personal innovativeness and device competence are the reasons for more job performance and satisfaction. They base their view on the Social Cognitive Theory of Bandura (1977), who defines self-efficacy as the extent to which a person believes in one's own ability to complete a task or reach a goal. In the context of CYOD, device self-efficacy is defined as 'the belief a certain device will enable a person to perform his or her task'. In this research satisfaction is defined as the combination of the perceived satisfaction (device self-efficacy) with the device preference. The device preference is measured by the number of people that would choose a certain CYOD device for a task (device competence). For an overview of the used constructs see the Employee Survey section on the left in Figure 4.

3.2 Employee Survey and Risk Assessment

For this research 126 respondents in four multinational organizations were surveyed. In order to observe corresponding findings across the companies, an overall study protocol was created (Yin, 2009). Besides the questionnaires on the use of devices and satisfaction, context interviews were held at the participating companies, to determine the (type of) CYOD policies. The four companies were:

Company 1 – a Dutch-headquartered Financial Accountancy firm, with 155.000 employees in 144 countries worldwide.

Company 2 – a 20.000 employee Media and learning multinational, headquartering in Finland.

Company 3 – a US-based multinational with business in Trading, Purchasing, Distributing grain and other Agricultural commodities, with 143.000 employees in 67 countries worldwide.

Company 4 – a Dutch-headquartered multinational producer of alcoholic beverages, with worldwide over 90.000 employees in 178 countries.

In the first section of the Employee Survey the respondents were asked which knowledge tasks they perform, and how their current device supports this task. Next, they were asked if they felt having a device of their own choice would improve their task performance, and if so, which device they would choose. Finally they were asked if they were willing to contribute in the device cost (see Figure3).

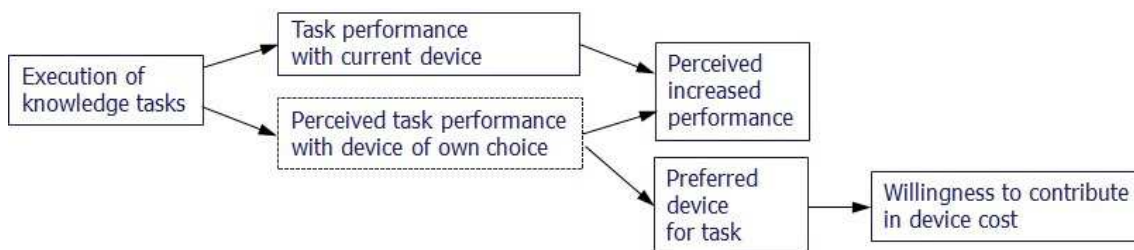


Figure 3. Employee Survey on tasks performance and preferred device

To determine the IT risks associated with the preferred CYOD devices, interactive Risk Assessments sessions were held with the IT Experts / Security Officers of the participating companies. For each

device the IT risk was determined and calculated, using two variables: the chance a threat can occur and the damage it will cause when it occurs. For the 7 before mentioned identified threats, in each participating company the IT expert or Security Officer evaluated the IT threats per type of device. The chance of occurrence and damage were rated on a 1 to 7 point Likert scale, meaning the highest risk for a specific threat for a device could be 49. The overall IT risk per device was determined by taking the average of all multiplications. The table below shows part of the used Device Risk Assessment sheet.

| Device Risk Assessment sheet | Chance of occurrence | | | | | | | | Damage | | | | | | | | |
|----------------------------------|----------------------|---|---------|---|---|-----------|---|-----|--------|----------|---|---------|---|---|-----------|-----|--|
| | Very Low | | Neutral | | | Very high | | | N/A | Very Low | | Neutral | | | Very high | | |
| Identified threats | 1 | 2 | 3 | 4 | 5 | 6 | 7 | N/A | 1 | 2 | 3 | 4 | 5 | 6 | 7 | N/A | |
| 1. Act of Human Error or Failure | | | | | | | | | | | | | | | | | |
| Device X | | | | | | | | | | | | | | | | | |

Table 2. Structure Risk Assessment sheet

In the second section of the Employee Survey the usefulness and satisfaction were investigated. The usefulness was determined based on the average of the first three constructs: Perceived usefulness, Perceived ease of use, and Quality of output. For the satisfaction, the fourth construct, Perceived satisfaction was combined (multiplied) with the score on preferred devices. The results of the Employee Survey were then combined with the Risk Assessment. This leads to the following analysis model:

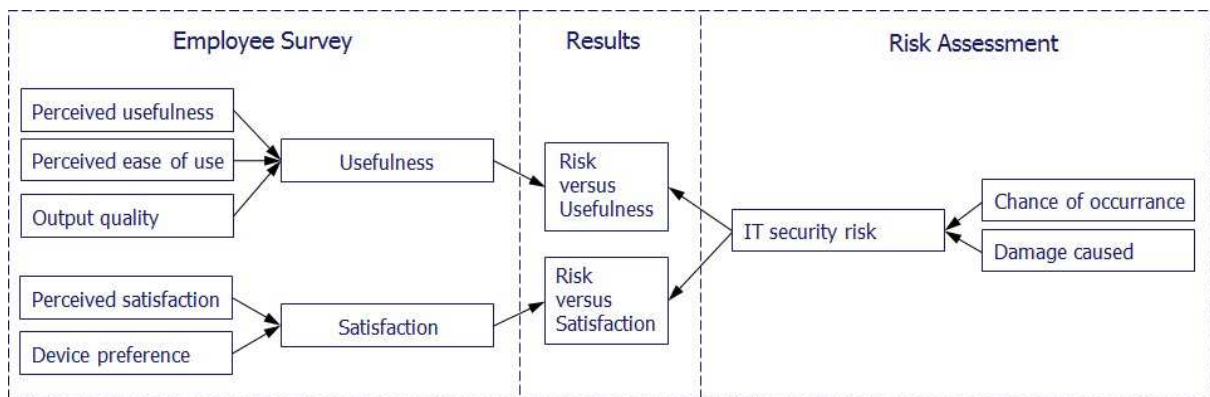


Figure 4. Analysis model

4 Research results

4.1 CYOD and related policies

In the context analysis, the CYOD policies were analyzed. All four companies have a CYOD policy in place, but the choices per device type differ. Also the use of own devices brought to the workplace differs per company. In most companies the use of own devices is restricted: Don't Bring Your Own Device (DBYOD), meaning personal devices can only be used on the guest network. Table 3 gives an overview of the CYOD policy and related policies at the participating companies.

| CYOD & BYOD policy | Notebooks | | | Mobile Phones | | | Tablets | | |
|--------------------|--------------|--------------|-------------|---------------|---------------------------|-------------|-------------|--------------|-------------------------|
| | Policy | CYOD options | BYOD access | Policy | CYOD options | BYOD access | Policy | CYOD options | BYOD access |
| Company 1 | CYOD : DBYOD | Win | Guest | CYOD : DBYOD | iOS +MDM | Guest | BYOD | None | Guest / Corp (iOS +MDM) |
| Company 2 | CYOD + BYOD | Win or iOS | Corporate | CYOD + BYOD | iOS (Guest) | Guest | CYOD + BYOD | iOS (Guest) | Guest |
| Company 3 | CYOD : DBYOD | Win | Guest | CYOD : DBYOD | iOS / Android + MDM | Guest | CYOD + BYOD | Win | Guest / Corp (iOS +MDM) |
| Company 4 | CYOD : DBYOD | Win | Guest | CYOD : DBYOD | Win / iOS / Android + MDM | Guest | BYOD | None | Guest / Corp (iOS +MDM) |

Table 3. CYOD and related policies

Table 3 shows that companies 1,3, and 4 have a CYOD & DBYOD policy in place for notebooks and mobile phones. For tablets they have a BYOD policy in place, at company 3 combined with the CYOD option. These three companies use a Mobile Device Management, MDM, tool to control the IT security risks. Company 2 has no MDM software in place, but is nevertheless allowing BYOD notebooks on the corporate network. The company has recognized this is an IT risk. For mobile phones and tablets there is both a CYOD and BYOD policy, but all devices are excluded from the corporate network. This makes company provided CYOD devices (as well as BYOD devices) relatively useless to perform business tasks on.

4.2 Tasks and performance

The overview of the tasks (knowledge actions) the respondents perform is in Appendix 1. As knowledge workers often perform more than one task, the total number of tasks is higher than the number (126) of respondents; in total 405 tasks were mentioned, meaning the average respondent performs a bit more than three (3,2) tasks. Analysis, Acquisition and Information search are the most performed knowledge tasks. These observations (multiple tasks per user/role and most frequent tasks/knowledge actions) are in line with the research results of Reinhardt et al. (2011).

When asked if the respondents believed their tasks could be performed well with the device they currently use, 53% of the respondents agree and 18% strongly agree that their current device supports the execution of their tasks well. Yet, when asked if they believe that having a device of their own choice, would increase their task performance, the response is as follows (Figure 5):

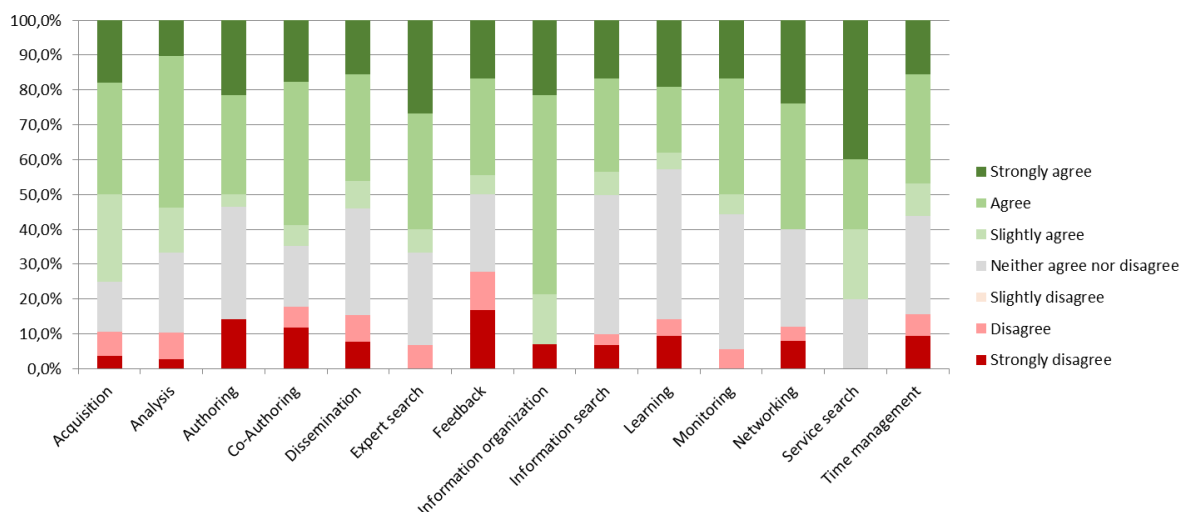


Figure 5. Perceived performance improvement per task with CYOD

Figure 5 shows that 52% of the respondents agree or strongly agree that having a device of their own choice would improve their task performance. This is an interesting outcome in the light of the first question, where over 70% of the respondents indicated to be able to perform their work well on their current device. The outcome does however fall in line with the research of Harris et al. (2012), who state that if employees were to choose their own hardware and software for work, they (strongly) agree that they would complete more tasks on time (49%), be more innovative (50%), and would be a happier employee (53%).

For each task, the respondents were asked whether they would rather use another device than the one they currently use, and if so, which device. The results of this analysis is in Appendix 2. The results show that a vast majority of the respondents would prefer another device, if given the choice in a CYOD environment. In general notebooks are preferred over desktops, provided they perform well enough. While only 2 respondents currently use an Apple notebook, Macbooks are preferred by most respondents over Windows notebooks. For the more mobile tasks light (and thin) notebooks or tablets

are preferred. Overall, the Apple iPad the is most preferred CYOD device, especially for reading and viewing data.

Appendix 3 gives an overview of the current devices in use for the tasks. The table in Appendix3 also contains the sum of the preferred devices that were mentioned. Figure 6 shows the relative spread of the current device use as well as the spread of the preferred devices in a CYOD environment.

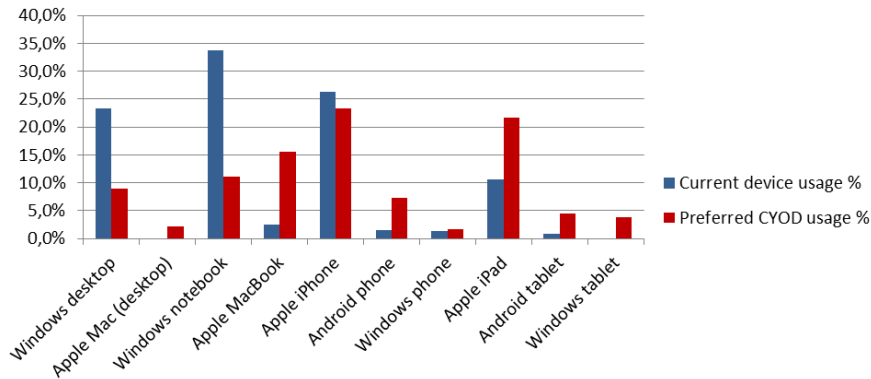


Figure 6. Current and preferred device usage

Figure 6 shows that in CYOD environment, there is less need for Windows desktops and notebooks, and more need for Apple notebooks, and in particular Apple iPad tablets. In general this means that, when implementing a CYOD policy that fits the preferences of the users, the number of operating systems and the number of different sorts of devices the ICT department has to manage, will rise.

Finally, the respondents were asked if they were willing to pay fully or partially for the device or their own preference. When it comes to paying, almost 75% (74,8%) of the respondents is not willing to contribute anything for the device of their choice. A group of around 15% is willing to pay up to 50% of the device cost. When the respondents are correct about the perceived improvement of their performance with the device of their choice, this would justify a CYOD policy (above a BYOD policy), as most of the employees are not willing to contribute personally to their improved business performance, but there is a lot of potential to gain.

4.3 Risk Assessment

The detailed results of the Risk Assessment can be found in Appendix 4. Though the Apple Mac desktop is not used in one of the case companies, the device was included in the Risk Assessment as it was one of the preferred CYOD devices. The risks that were determined are the net risks of the devices, meaning that the risk degree already includes a proper security policy with technical controls in place. The overall IT risk of the devices is determined by calculating the average of the outcomes of all participating companies. This result is shown in Figure 7.

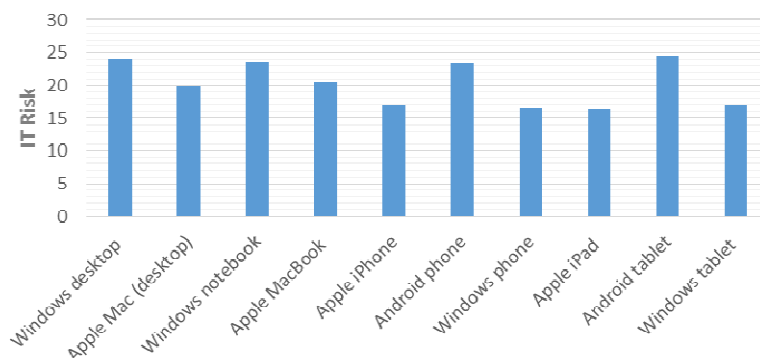


Figure 7. IT Risk for devices

Figure 7 shows that Windows desktops and notebooks, and Android phones and tablets, are the devices with the highest IT risks. Windows phones and tablets, and Apple devices in general, are the devices with the lowest IT security risks.

4.4 Usefulness and satisfaction

In the second section of the Employee Survey, the respondents were asked to score devices on Perceived usefulness, Perceived ease of use, and Output quality for each knowledge task. The results of this analysis is in Appendix 5. Overall, Windows notebooks score well on Perceived usefulness, Perceived ease of use, and Output quality. The iPad is less suitable for tasks e.g. Authoring and Analysis, but more suitable for reading and viewing tasks. Both laptops and tablets are suitable for Information search; tasks where mobile phones (iPhone and Windows phone) score lower. Finally, the respondents were asked to score the device of their own choice on Perceived satisfaction. Figure 8 shows the overall results of the Usability and Perceived satisfaction outcomes for with the different devices.

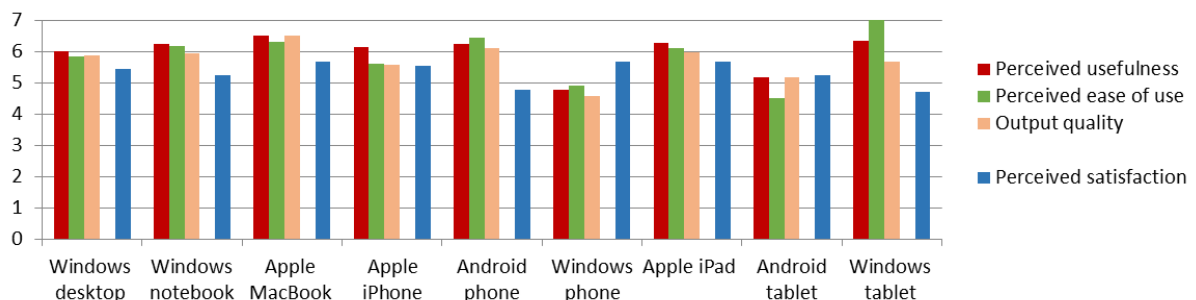


Figure 8. Usefulness and perceived satisfaction per device

In general, the perceived satisfaction scores lower than usability (the three constructs), except for the Windows phone. Possibly this is because the respondent were cautious of being over-optimistic.

4.5 IT Risk versus usefulness and satisfaction

When the Risk Assessment result is plotted against the Usefulness, being the average of the constructs: Perceived usefulness, Perceived ease of use and Output quality, the following picture appears.

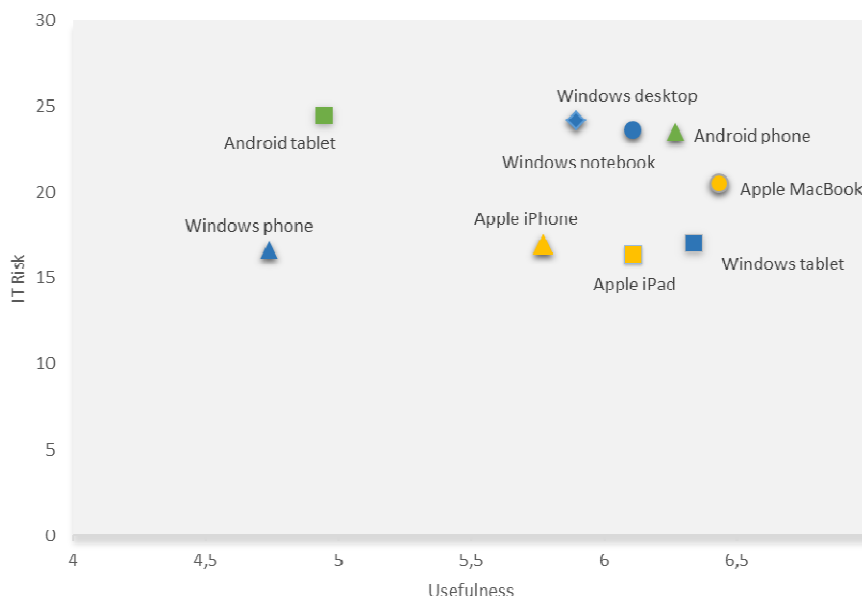


Figure 9. Usefulness versus IT Risk

Figure 9 shows the Android tablet and Windows phone score lowest on usefulness, while the Apple MacBook and Windows tablet score highest. From an IT security point of view, Windows desktops and notebooks and Android tablets and phones score worst.

When the IT Risk is plotted against the Satisfaction, being the combination (multiplication) of the Perceived satisfaction with the number of choosers of a preferred device, a quite different picture appears.

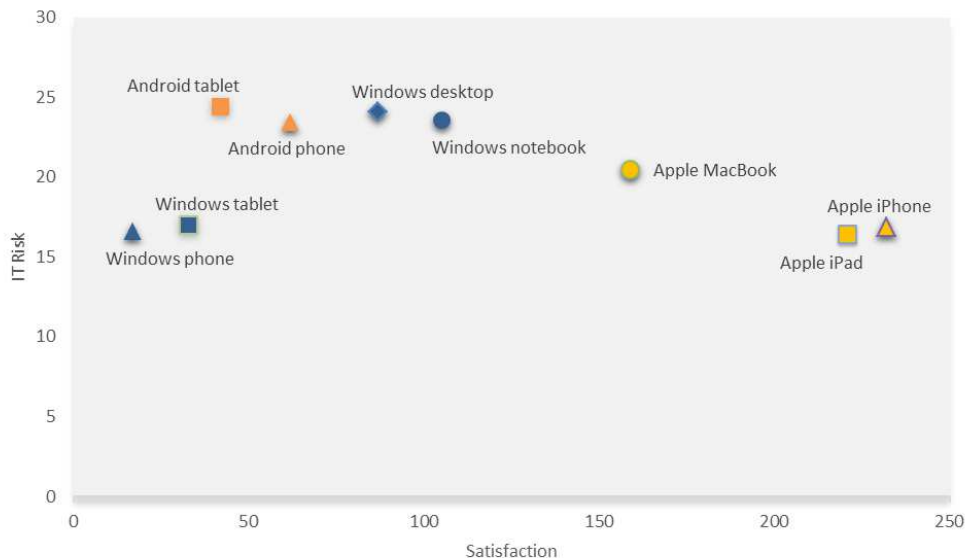


Figure 10. Satisfaction versus IT Risk

Figure 10 shows that Apple devices score by far best when it comes to satisfaction (preferred device and perceived satisfaction). Windows desktops and notebooks are somewhere in the middle, while Android and Windows phones and tables are at the bottom of the preference list. From an IT security point of view the preferred CYOD devices are less vulnerable than the Windows devices, that are often currently in use. This leads to an interesting conclusion: enabling employees to improve their task performance, whilst experiencing a higher job satisfaction, by giving them the opportunity to use a device or their own choice in a CYOD environment, does not increase, but instead reduces, the overall average IT security risks. A precondition for the above situation is that the proper security policies with technical controls are in place. This means that the implementation of a CYOD policy (with more Apple devices) does not raise the IT risk level, but it does mean the management of more platforms and software.

5 Discussion, conclusions and future research

5.1 Discussion

As mentioned in the introduction, research on Choose Your Own Device (CYOD) policies in the area of an implementation of the New Way of Working is scarce. Comparable literature on NWOW and CYOD can hardly be found, if any. This research on IT security risks versus usability and device satisfaction, in a NWOW and CYOD environment, is possibly one of the first steps in this area. Some critical notes are however at its place.

Having four companies with 126 respondents is reasonable, but the respondents were not evenly distributed across the organizations. This made intra-company comparisons unreliable if not impossible, and has the risk of over-emphasizing company-related viewpoints.

The perceived satisfaction and number of preferred devices for a task are subjective user-perceptions. It may well be that an Apple iPad is in reality not the best device for the given task, even if respond-

ents believe it is. This effect (likability versus reality) has not been measured, but is realistic in both this research as in daily business practice. This may mean that, though in reality a Windows tablet could be more useful for executing a task than an Apple iPad, most users would still prefer an iPad, when given the choice, to perform their task on.

Having the IT Experts and Security Officers of four multinationals available for the Risk Assessment is good, but estimating risks remains a subjective and human exercise. The results should therefore be seen as a first indication of the possible effects of CYOD on job performance and employee satisfaction.

5.2 Conclusions and future research

Organizations struggle with the phenomenon of employees using consumer devices for business purposes. In an optimal situation the use of these personalized mobile devices would be beneficial for both the employee and organization, rendering higher employee satisfaction with higher performance on task execution. The question is how this optimum can be reached. Having researched the IT security risks against the effects and possible gains of a CYOD policy, this study shows that:

- Though over 70% of the respondents agree they can perform their tasks well with their current device, a majority (52%) of the respondents (strongly) agrees, having the ability to use a device of their own, will increase their task performance.
- The vast majority of employees, almost 75%, is however not willing to contribute to the costs of personal devices. Combined with the first conclusion, this implies that a CYOD policy is to be preferred over a BYOD policy, and can be beneficial for the organization.
- Introducing a CYOD environment in an organization will lead to a shift in the types of devices used. Desktops are likely to be replaced by (powerful) notebooks, preferably in combination with optional large monitors, and where suitable for the task, tablets will be used instead of notebooks.
- The introduction of the CYOD environment will lead to the mandatory management of more platforms and software. Besides Windows devices, Apple devices and the use of (iOS) apps will need to be fully supported by the corporate IT strategy.
- Under the precondition that the security policies with technical controls are in place, the introduction of a CYOD policy does not necessarily increase the level of IT security risk. The average net IT risk may even decrease when introducing CYOD, e.g. in this research with the preferred Apple devices.

Enabling employees to improve their task performance whilst experiencing a higher job satisfaction, by giving them the opportunity to use a device of their own choice and preference, in a CYOD environment, does not by definition increase the overall average IT security risks. Organizations that know which devices employees need to best perform their tasks, can balance out the business risk requirements and meet the employee expectations to maximize employee satisfaction without giving up on corporate data protection. In doing so, the consequence will be the management of more platforms and operating systems in a controlled CYOD environment.

This research is only a first step towards a future of effective CYOD policies in a NWOW environment. There will always be more information to explore and describe. For instance: the aspect of the usefulness of software in combination with (preferred) hardware was not researched in this study, but is certainly an aspect worth investigating in future studies of CYOD in a NWOW environment. Also the cost of a CYOD program against the possible business gain could be a field of future study, as well as the actual performance gain from implementing CYOD in real business practice. The results of this study should therefore be used with care, as more future research should support these first findings, and add more insights.

References

- Ajzen, I. & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Baane R., Houtkamp P., & Knotter M. (2010). *The new world of work unravelled – Het nieuwe werken ontrafeld – over Bricks, Bytes & Behavior*. 1-168. Koninklijke Van Gorcum BV. ISBN 9789023245858.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review* (84:2), pp. 191–215.
- Baskerville, R. (1993). Information systems security design methods: implications for information systems development. *ACM Computing Surveys*, 25(4), 375–414. <http://doi.org/10.1145/162124.162127>
- Bijl, D.W. (2011). *Journey towards the New Way of Working - creating sustainable performance and joy at work*. Par CC. ISBN: 978-94-90528-00-3
- Bødker, S., & Christiansen, E. (2002). *Lost and Found in Flexibility*. University of Aarhus, Center for New Ways of Working. Retrieved from [http://pure.au.dk/portal/en/publications/lost-and-found-in-flexibility\(cd3361c0-983b-11da-bee9-02004c4f4f50\).html](http://pure.au.dk/portal/en/publications/lost-and-found-in-flexibility(cd3361c0-983b-11da-bee9-02004c4f4f50).html)
- Citrix. (2013). *Best practices to make BYOD simple and secure*. Citrix Systems.
- Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. San Francisco, CA: Jossey-Bass.
- Csikszentmihalyi, M. (1988). The flow experience and human psychology. In M. Csikszentmihalyi & I. S. Csikszentmihalyi (Eds.) (1988). *Optimal Experience: Psychological studies of flow in consciousness* (pp. 15-35). New York: Cambridge University Press.
- Csikszentmihalyi, M. (1990). *Flow: the psychology of optimal experience*. New York: Harpers Perennial.
- Davenport, T.H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Boston: Harvard Business School Press.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-342.
- Finneran, C. M., & Zhang, P. (2002). *The Challenges of Studying Flow within a Computer-Mediated Environment*. Paper presented at the Americas Conference on Information Systems, Dallas, TX.
- Gajar, P. K., Ghosh, A., & Rai, S. (2013). Bring Your Own Device (BYOD): Security risks and mitigating strategies. *Journal of Global Research in Computer Science*, 4(4), 62–70.
- Gartner. (2015). *Identity and Access management (IAM)*. Retrieved from <http://www.gartner.com/it-glossary/identity-and-access-management-iam/>
- Ghani, J. (1991). Flow in human computer interactions: test of a model. In J. Carey (Ed.), *Human Factors in Information Systems: Emerging Theoretical Bases*. New Jersey: Ablex Publishing Corp.
- Giddens, L., & Tripp, J. (2014). It's My Tool , I Know How to Use It : A Theory of the Impact of BYOD on Device Competence and Job Satisfaction. *Proceedings of the Twentieth Americas Conference on Information Systems*, pp. 1–8.
- Gillett, F. E. (2012). *Tablets Will Rule The Future Personal Computing Landscape*. Forrester Research.
- Govcert (2009). *Security of mobile devices and data carriers – Dutch title: Beveiliging van mobiele apparatuur en datadragers*. National Cyber Security Center. The Hague. Retrieved from: <https://www.ncsc.nl/dienstverlening/expertise-advies/kennisdeling/whitepapers/whitepaper-beveiliging-van-mobiele-apparatuur-en-datadragers.html>
- Greene, C., & Myerson, J. (2011). Space for thought: designing for knowledge workers. *Facilities*, 29(1/2), 19-30.
- Hammer, M., & Champy, J.A. (1993) *Reengineering the Corporation: A Manifesto for Business Revolution*. Ch. 4 & 5. HarperCollins Publishers Inc. ISBN 0-06-662112-7.
- Harris, J., Ives, B., & Junglas, I. (2012). IT Consumerization: When gadgets turn into enterprise IT tools. *MIS Quarterly Executive*, 11(3), 99–112.
- Holtznider, B., and Jaffe, B. D. (2012) *IT Manager's Handbook: Getting Your New Job Done*. Morgan Kaufmann.

- Ingalsbe, J. a, Shoemaker, D., & Mead, N. R. (2011). Threat Modeling the Cloud Computing, Mobile Device Toting, Consumerized Enterprise – an overview of considerations. AMCIS 2011 Proceedings, 1–6.
- Johns, T., & Gratton, L. (2013). The Third Wave Of Virtual Work. Harvard Business Review, January-February 2013, pp. 66-73.
- Jones, P., & Jordan, J. (1998). Knowledge orientations and team effectiveness. International Journal of Technology Management, 16, 152-161.
- Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. Internet and Higher Education, 8(1), 13–24. <http://doi.org/10.1016/j.iheduc.2004.12.001>
- Kok, A. de, Bellefroid, B.E.W. & Helms, R.W. (2013) Knowledge Sharing and Channel Choice: Effects of the New Way of Working. Proceedings of the 14th European Conference on Knowledge Management, ECKM 2013.
- Kok, A. de, Koops, J. & Helms, R.W. (2014) Accessing the New Way of Working: Bricks, Bytes and Behaviour. Proceedings of the 18th Pacific Asia Conference on Information Systems, PACIS 2014.
- Morrow, B. (2012). BYOD security challenges: Control and protect your most sensitive data. Network Security, 2012(12), 5–8. doi:10.1016/S1353-4858(12)70111-3
- Moschella, D., Neal, D., Opperman, P., and Taylor, J. 2004. The “Consumerization” of Information Technology. El Segundo, California, USA: CSC.
- Niehaves, B., Köffer, S., & Ortbach, K. (2012) IT Consumerization – A Theory and Practice Review, AMCIS, paper 18.
- Niehaves, B., Köffer, S., & Ortbach, K. (2013) The Effect of Private IT Use on Work Performance - Towards an IT Consumerization Theory. International Conference on Wirtschaftsinformatik: Leipzig.
- Peltier, T. R. (2005). Information Security Risk Analysis. Auerbach Publications, CRC Press, Taylor & Francis Group. ISBN 0-8493-3346-6.
- Reinhardt, W., Schmidt, B., & Sloep, P. (2011). Knowledge worker roles and actions - Results of two empirical studies. Knowledge and Process Management, 18(3), 150–174. <http://doi.org/10.1002/kpm>
- Rouse. (2007). Authentication.
Retrieved from <http://searchsecurity.techtarget.com/definition/authentication>
- Ruch, T., & Gregory, R. (2014). Consumerization of It – Where Is the Theory? Proceedings of the 18th Pacific Asia Conference on Information Systems, PACIS 2014.
- Venkatesh, V., & Davis, F. (1996). A Model of the Antecedents of Perceived Ease of Use: Development and Test. Decision Sciences, 27(3), 451-481.
- Venkatesh, V., & Davis, F. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. Management Science, 46(2), 186–204.
- Waber, B., Magnolfi, J., & Lindsay, G. (2014). Workspaces that move people. Harvard Business Review, Oktober 2014, pp. 68–77.
- Whitman, M. E. (2003). Enemy at the Gate: Threats to Information Security. Communications of the ACM, 46(8), 91–95. Retrieved from <http://portal.acm.org/citation.cfm?id=859675>
- Yin, R.K. (2009). Case Study Research. Thousand Oaks , California, USA, Sage Publications.

Appendix 1 Knowledge tasks

Typology of knowledge actions / tasks of Reinhardt et al. (2011) and their description. In this research the task 'Time management' was added to the list of Reinhardt, because managing time was identified as an important part of the tasks of a large number of the respondents (e.g. in consulting work).

| Knowledge action | Description |
|---------------------------------|--|
| Acquisition | The gathering of information with the goal of developing skills or project or obtaining an asset. |
| Analyze | The examining or thinking about something carefully, in order to understand. |
| Authoring | The creation of textual and medial content using software systems, for example word processing systems/ presentation systems. |
| Co-authoring | The collaborative creation of textual and medial content using software applications, for example, word processing systems/ presentation software. |
| Dissemination | The spreading of information or information objects, often work results. |
| Expert search | The retrieval of an expert to discuss and solve a specific problem. |
| Feedback | The assessment of a proposition or an information object. |
| Information organization | The personal or organizational management of information collection. |
| Information Search | The looking up of information on a specific topic and in a specific form. Often we search using the folder structure of a file system or we search using an information retrieval service. |
| Learning | The acquiring process of new knowledge, skills or understanding during the execution of work or based on formalized learning material. |
| Monitoring | Keeping oneself or the organization up-to date about selected topics, for example, based on different electronic information resources. |
| Networking | The interaction with other people and organizations to exchange information and develop contacts. |
| Service search | The retrieval of specialized web services that offer specific functions. |
| Time management | The planning, recording and invoicing of time spend on work activities. |

Table 4. *Typology of knowledge actions / tasks (Reinhardt et al., 2011)*

Overall number of tasks performed by the respondents.

| Task | #Respondents | % | Task | #Respondents | % |
|---------------|--------------|-------|--------------------------|--------------|-------|
| Acquisition | 45 | 11,1% | Information organization | 22 | 5,4% |
| Analysis | 53 | 13,1% | Information Search | 41 | 10,1% |
| Authoring | 40 | 9,9% | Learning | 24 | 5,9% |
| Co-authoring | 18 | 4,4% | Monitoring | 23 | 5,7% |
| Dissemination | 19 | 4,7% | Networking | 27 | 6,7% |
| Expert search | 23 | 5,7% | Service search | 5 | 1,2% |
| Feedback | 25 | 6,2% | Time management | 40 | 9,9% |

Table 5. *Number of tasks (knowledge actions) performed by respondents*

Appendix 2 Preferred CYOD devices

- For Acquisition, 28 (out of 35) respondents indicate they would prefer another device for their task. The type of preferred device varies. Respondents with a desktop computer prefer a notebook (Apple or Windows). Some respondents using notebooks or iPhones prefer larger screens than an iPhone, but smaller screens than their notebook. Performing acquisition tasks on an Android phone is perceived by them as useful and easy, though they see the output quality as low. Overall, the device that is perceived as best for Acquisition tasks is an Apple notebook.
- For Analysis a fast computer (e.g. a desktop) is often preferred over the current devices (Windows notebooks). A thin laptop (Windows or Apple) or a tablet is perceived as useful for traveling and out-of-office work. Performing Analysis on an Apple notebook is questioned as respondents didn't find the device useful and easy to use, and the output quality not high. The Windows notebook scores low on perceived usefulness and perceived ease of use, though the output quality scores high. Overall, the Windows desktop scores best for performing analysis tasks.
- For Authoring and Co-authoring, many respondents consider a laptop as the best device for their task. Also a tablet (iPad with supporting apps) is preferred, because it is easier to carry. Some respondents prefer a thin and light laptop (Windows or MacBook Air). For authoring tasks both Windows notebooks and Apple MacBooks score high. Although mobile devices are also used for authoring, those devices are perceived as less suitable than laptops.
- For Dissemination of information respondents currently use desktop devices, but prefer an Apple or Windows notebook. Respondents already using a notebook prefer a faster and thinner laptop. Also a tablet (iPad) was indicated as a (more) useful device for this task.
- For Expert search, 20 (out of 23) respondents would rather use another device. Different devices are mentioned, such as Apple and Windows notebooks. Also a newer version of the iPhone device is preferred, with the bigger screen for mobile apps such as LinkedIn.
- For Feedback, 20 (of 25) respondents indicate to prefer another device. Suggestions include a thinner and smaller Windows notebook or Apple MacBook Air. Also iPads and Android tablets were suggested as useful.
- For Information organization a light laptop (e.g. MacBook Air) is preferred, or a notebook instead of a desktop. Also iPads are mentioned several times. One respondent (now using a notebook) replied; "A windows notebook is fine, but I do not have the software to manage disparate flows of data information. A device with such software would be my preferred device."
- For Information search, some respondents emphasized that the devices hardly matters, provided that it has a good way of conveying the information. It is the search software that matters to them. Though the hardware is said to not matter, still 35 out of the 41 respondents prefer other devices than they currently use. All types of other devices are mentioned: Apple MacBooks, Windows notebooks, larger Phones/iPads, Android tablets, and Windows tablets.
- For Learning, all (24) respondents indicate to rather use other devices. They prefer an Apple MacBook or Windows notebook over their current desktop computer. Also a tablet (iOS or Android) is mentioned as a preferred device.
- For Monitoring, 19 (of 23) respondents prefer other devices. They differ from Apple MacBooks to Windows notebooks or tablets instead of mobile phones due to the screen size.
- For Networking, 26 (of 27) respondents rather use another device than they currently use. One respondent rather uses a Blackberry phone, another rather uses an Android (Samsung) smart phone or tablet instead of iPhone. A tablet is mentioned several times, including iPad, Android tab and a Windows tablet.
- For Service search, all (5) respondents preferred another device for searching services. The only mentioned devices are Apple MacBook and Windows notebook.
- For Time management, 37 (of 40) respondents prefer another device. Tablets are in favor (iPad, Android, or Windows).

Appendix 3 – Current and preferred use of devices

| <i>Devices currently in use</i> | <i>Windows desktop</i> | <i>Apple Mac (desktop)</i> | <i>Windows notebook</i> | <i>Apple MacBook</i> | <i>Apple iPhone</i> | <i>Android phone</i> | <i>Windows phone</i> | <i>Apple iPad</i> | <i>Android tablet</i> | <i>Windows tablet</i> |
|-----------------------------------|------------------------|----------------------------|-------------------------|----------------------|---------------------|----------------------|----------------------|-------------------|-----------------------|-----------------------|
| Tasks | | | | | | | | | | |
| <i>Acquisition</i> | 10 | 0 | 23 | 2 | 29 | 3 | 0 | 5 | 0 | 0 |
| <i>Analysis</i> | 22 | 0 | 24 | 2 | 26 | 0 | 1 | 9 | 3 | 0 |
| <i>Authoring</i> | 13 | 0 | 24 | 2 | 7 | 0 | 1 | 5 | 0 | 0 |
| <i>Co-authoring</i> | 8 | 0 | 10 | 1 | 0 | 0 | 1 | 2 | 0 | 0 |
| <i>Dissemination</i> | 6 | 0 | 12 | 0 | 8 | 0 | 1 | 6 | 0 | 0 |
| <i>Expert search</i> | 9 | 0 | 12 | 2 | 11 | 1 | 0 | 6 | 0 | 0 |
| <i>Feedback</i> | 10 | 0 | 13 | 0 | 10 | 0 | 1 | 2 | 0 | 0 |
| <i>Information organization</i> | 9 | 0 | 10 | 1 | 6 | 0 | 0 | 3 | 0 | 0 |
| <i>Information Search</i> | 16 | 0 | 21 | 1 | 28 | 1 | 1 | 10 | 1 | 0 |
| <i>Learning</i> | 10 | 0 | 13 | 1 | 5 | 2 | 1 | 3 | 0 | 0 |
| <i>Monitoring</i> | 8 | 0 | 13 | 1 | 7 | 1 | 0 | 4 | 0 | 0 |
| <i>Total current device usage</i> | 121 | 0 | 175 | 13 | 137 | 8 | 7 | 55 | 4 | 0 |
| <i>Current device usage %</i> | 23,3% | 0,0% | 33,7% | 2,5% | 26,3% | 1,5% | 1,3% | 10,6% | 0,8% | 0,0% |
| <i>Preferred CYOD usage</i> | 16 | 4 | 20 | 28 | 42 | 13 | 3 | 39 | 8 | 7 |
| <i>Preferred CYOD usage %</i> | 8,9% | 2,2% | 11,1% | 15,6% | 23,3% | 7,2% | 1,7% | 21,7% | 4,4% | 3,9% |

The current device usage % is the relative spread of the current devices in use, in relation to the total number of current devices.

The preferred CYOD usage % is the relative spread of the number of preferred devices, in relation to the total number of preferred devices.

Appendix 5 – Evaluation of usefulness of devices

| <i>Device in use</i> | <i>Windows desktop</i> | <i>Windows notebook</i> | <i>Apple MacBook</i> | <i>Apple iPhone</i> | <i>Android phone</i> | <i>Windows phone</i> | <i>Apple iPad</i> | <i>Android tablet</i> | <i>Windows tablet</i> |
|---------------------------------|------------------------|-------------------------|----------------------|---------------------|----------------------|----------------------|-------------------|-----------------------|-----------------------|
| <i>Task</i> | | | | | | | | | |
| Acquisition | | | | | | | | | |
| <i>Usefulness</i> | 6 | 5,9 | 6,5 | 7 | 5,3 | - | 5,8 | - | - |
| <i>Ease of use</i> | 5,7 | 5,6 | 6,5 | 6,7 | 5,3 | - | 5,4 | - | - |
| <i>Output quality</i> | 5,8 | 5,5 | 7 | 4,7 | 5,3 | - | 6 | - | - |
| Analysis | | | | | | | | | |
| <i>Usefulness</i> | 6,2 | 5,7 | 5 | - | 5,6 | 2 | 5,8 | 4,3 | - |
| <i>Ease of use</i> | 6,2 | 5,3 | 4,5 | - | 5,6 | 1 | 5,6 | 4 | - |
| <i>Output quality</i> | 6,1 | 6,3 | 4,5 | - | 5,6 | 1 | 5,7 | 4,3 | - |
| Authoring | | | | | | | | | |
| <i>Usefulness</i> | 6,2 | 6,4 | 6,5 | - | 5,8 | 5 | 5,6 | - | - |
| <i>Ease of use</i> | 5,8 | 6,2 | 6,5 | - | 5,6 | 5 | 5 | - | - |
| <i>Output quality</i> | 5,7 | 6,2 | 6,5 | - | 5,1 | 4 | 4,8 | - | - |
| Co-authoring | | | | | | | | | |
| <i>Usefulness</i> | 6,1 | 6,5 | 7 | - | - | 4 | 6,5 | - | - |
| <i>Ease of use</i> | 5,8 | 6,4 | 7 | - | - | 5 | 6,5 | - | - |
| <i>Output quality</i> | 6 | 6,2 | 7 | - | - | 4 | 6,5 | - | - |
| Dissemination | | | | | | | | | |
| <i>Usefulness</i> | 6,2 | 6,3 | - | - | 5,3 | 5 | 6,2 | - | - |
| <i>Ease of use</i> | 6 | 6,5 | - | - | 5,8 | 5 | 6,2 | - | - |
| <i>Output quality</i> | 6 | 6,3 | - | - | 5,5 | 5 | 6,2 | - | - |
| Expert search | | | | | | | | | |
| <i>Usefulness</i> | 5,9 | 6,3 | 6,5 | 7 | 5,5 | - | 6,2 | - | 5 |
| <i>Ease of use</i> | 6 | 6,2 | 6 | 7 | 5,2 | - | 6,2 | - | - |
| <i>Output quality</i> | 6 | 5,8 | 6 | 7 | 4,9 | - | 5,7 | - | 3 |
| Feedback | | | | | | | | | |
| <i>Usefulness</i> | 5,9 | 6 | - | 5 | 5,9 | 5 | 6,5 | - | - |
| <i>Ease of use</i> | 5,9 | 6,1 | - | 6 | 5,8 | 6 | 6,5 | - | - |
| <i>Output quality</i> | 6 | 5,1 | - | 6 | 5,8 | 5 | 6,5 | - | - |
| Information organization | | | | | | | | | |
| <i>Usefulness</i> | 5,8 | 5,9 | 6 | 5,5 | 5,5 | - | 6 | - | - |
| <i>Ease of use</i> | 5,7 | 5,6 | 6 | 6 | 5,8 | - | 6 | - | - |
| <i>Output quality</i> | 5,7 | 5,4 | 6 | 6 | 5,5 | - | 6,7 | - | - |

Appendix 5 - Evaluation of usefulness of devices (continued)

| Device | Windows desktop | Windows notebook | Apple MacBook | Apple iPhone | Android phone | Windows phone | Apple iPad | Android tablet | Windows tablet |
|---------------------------|-----------------|------------------|---------------|--------------|---------------|---------------|------------|----------------|----------------|
| Task | | | | | | | | | |
| Information Search | | | | | | | | | |
| Usefulness | 5,9 | 6,4 | 7 | 5 | 5,5 | 5 | 6 | 6 | 7 |
| Ease of use | 5,5 | 6,3 | 6 | 6 | 5,4 | 5 | 5,9 | 5 | 7 |
| Output quality | 5,6 | 6 | 7 | 6 | 5,7 | 5 | 6,2 | 6 | 7 |
| Learning | | | | | | | | | |
| Usefulness | 5,9 | 6,3 | 7 | 5,5 | 6 | 5 | 6,7 | - | - |
| Ease of use | 5,8 | 6,3 | 6 | 6 | 6 | 5 | 7 | - | - |
| Output quality | 5,8 | 6,3 | 7 | 6 | 6 | 5 | 6 | - | - |
| Monitoring | | | | | | | | | |
| Usefulness | 6 | 6,3 | 7 | 7 | 6 | - | 6,5 | - | - |
| Ease of use | 5,8 | 6,2 | 7 | 7 | 5,9 | - | 6,3 | - | - |
| Output quality | 5,8 | 6,1 | 7 | 7 | 5,9 | - | 6 | - | - |
| Networking | | | | | | | | | |
| Usefulness | 5,8 | 6,3 | 7 | 6 | 6,2 | 5 | 6,5 | - | 7 |
| Ease of use | 5,6 | 6,3 | 7 | 6 | 6,1 | 5 | 6,5 | - | 7 |
| Output quality | 5,8 | 6,3 | 7 | 6 | 6,1 | 5 | 6,3 | - | 7 |
| Service search | | | | | | | | | |
| Usefulness | - | 6,5 | 7 | - | 5,7 | - | 7 | - | - |
| Ease of use | - | 6,5 | 7 | - | 4,7 | - | 6 | - | - |
| Output quality | - | 6,5 | 7 | - | 5,3 | - | 5 | - | - |
| Time management | | | | | | | | | |
| Usefulness | 6,1 | 6,3 | 6 | - | 6,1 | 7 | 6,5 | - | - |
| Ease of use | 6 | 6,3 | 6 | - | 5,7 | 7 | 6,3 | - | - |
| Output quality | 5,8 | 6,1 | 6 | - | 5,7 | 7 | 6,2 | - | - |
| Overall | | | | | | | | | |
| Usefulness | 6,0 | 6,2 | 6,5 | 6,0 | 5,7 | 4,8 | 6,3 | 5,2 | 6,3 |
| Ease of use | 5,8 | 6,1 | 6,3 | 6,3 | 5,6 | 4,9 | 6,1 | 4,5 | 7,0 |
| Output quality | 5,9 | 6,0 | 6,5 | 6,1 | 5,6 | 4,6 | 6,0 | 5,2 | 5,7 |

NETWORKING WEBSITES USAGE: ALL STUDENTS MULTITASK, DO ALL MANAGE COGNITIVE LOAD? A COMPARATIVE STUDY

Research in Progress

Rouis, Sana, ESSCA School of Management, France, sana.rouis@essca.fr

Ayadi, Faouzi, College of Economics and Administrative Sciences, Al Imam Mohammad Ibn Saud Islamic University (IMSIU) – Riyadh, Kingdom of Saudi Arabia, fayadi2000@yahoo.fr

Zarbout, Rim, University of Management and Economic Sciences, Sfax, Tunisia, rim.zarbout@yahoo.fr

Abstract

Social networking has been expanding widely for almost two decades. All students use social networking platforms and usage is more intense and complex to impact a whole era social life. Based on cognitive load theory and dual task approach, we measure the mediating effect of social networking sites usage on the impact of individuals' characteristics on their academic achievement. Research model proposed presumes that basic values, personality traits and social status have an effect on Facebook usage and cognitive absorption which in turn have an effect on academic results. Research looks at Facebook usage for recreation into multitasking context and assesses moderating role of polychronicity capacity. Group comparison is used to analyze countries and gender differences.

Keywords: Facebook usage, cognitive absorption, basic values, personality traits, social status, polychronicity, academic achievement.

1 Introduction

For many years, there has been a widespread concern on students' performance. Many studies explored the impact of external aspects such as: students' socio-economic backgrounds (Tajalli and Opheim, 2004) and available opportunities (Boscardin et al., 2005). Others focused on the amount of effort shown by students in terms of their participation, discussion and interaction, either in class or during online sessions (Davies and Graff, 2005). Yet, others studies were concerned with educational methodologies such as the different methods used to assess students work, as well as the multiplicity of the information sources effect on students' performance (Greer, 2001; Funkhouser, 2003). All in all however, few studies looked at the impact of whether and how student's usage of the technology influences their achievement.

Based on cognitive load theory, past research focused on different ways information is presented to learners, bearing in mind the amount of distraction experienced by students as a result of being simultaneously attracted by different learning sources. In the present study we attempt to use the cognitive load theory to investigate the impact of the use of technology, outside the classroom, on students' performance. Special emphasis will be given to situations where, for example students are logged onto a social networking site while simultaneously doing their homework. Particular attention will be given to the usage of social media (SM) technologies and the cognitive load effect that these have on students. Such use is primarily being considered as a second task performed by students in order to have a better conditioning to improve their level of academic work.

A known fact is that many students spend most of their free time surfing the social websites. The advantages of this new social activity include a higher level of integration between teenagers and students, new friends are easy to find, as well as the continuous exchange of information and knowledge between students. On the other hand, unlimited exposure to these media is known to have an overall negative effect on students' performance. A browse through the pages of the more popular social networking sites will quickly reveal that many students constantly express their concern regarding their inability to concentrate on their academic work because they heavily distracted by these media.

This trend has been confirmed in a research study by Karpinski and Duberstein (2009), who concluded that in general, students using Facebook score lower grades than students who do not have access to such communities. While the former only spend an average of one to five hours a week to study at home, the average weekly study-time spent by non-users exceeds eleven hours. This situation can be attributed to an inappropriate direction of the students' attention which in turn leads to distraction, due to simultaneous multitask performance (Paas et al., 2003a; Paas et al. 2003b; DeLeeuw and Mayer, 2008).

On another hand, according to Karpinski and Duberstein, (2009), 79% of Facebook users (students) actually claim that spending time on social network sites does not interfere with their studies. Community members typically spend long hours interacting with each other, uploading videos and pictures, sharing web links, checking their inbox messages and sending their replies. Nevertheless, the number of active users of Facebook is steadily increasing, notwithstanding the adverse effects that this popular site may have on students' overall academic performance.

Indeed, educational authorities are rightfully concerned about how these media are affecting their students. On the other side, both parents and instructors need to understand the keen interest that students seem to develop for these technological wonders. Moreover, they should also learn how to react to this new trend. In this study, we question whether Facebook usage and cognitive absorption have negative effect on academic achievement on all students profiles combined?

The objectives of this research in progress paper would be to:

- Draw a research model that assesses the effect of Facebook usage on cognitive absorption and academic achievement and

- Develop the research hypotheses that account of the role of students personal characters in determining Facebook usage and cognitive absorption level.

2 Theoretical Background

Cognitive Load Theory (CLT), which aligns dual task approach, memory load, the distribution of the cognitive resources and split attention paradigm, is the underlying theory for this study.

2.1 Cognitive Load Theory (CLT)

As an instructional theory, CLT focuses on load generated by a dual tasks carried out in the same time and assumes that individuals' working memory can only process limited number of items at the same time. Research on cognitive load has been the object of many experimental and on site studies conducted with the aim to forecast best instructional context to stipulate students to have better achievements (Brünken et al., 2003), but also, CLT argues that individual working memory load may be affected diversely depending on the sources for cognitive load (Brünken et al., 2003). It can result from the inherent nature of the learning materials presented to the students, so called intrinsic cognitive load; the way the content is presented to the learners (extraneous cognitive load) or the quantity of cognitive resources people are willing to devote for the learning task, so called germane load.

Besides, Paas (1992) distinguishes two components for the cognitive load that are mental load and mental effort. In fact, mental load refers to the load induced by instructional parameters such as task structure and sequence of the information and the learner's interest to the subject. It makes a cohesive presentation of the intrinsic, extraneous and germane load. Mental effort, however, refers to the amount of capacity that is allocated to understand the learning materials and get knowledge out of it; it makes a slight difference with the germane cognitive load. While the germane load relates to the connection the learner is having or bearing on the subject taught, the mental effort relays more to the endeavor learners are willing to bring in in order to achieve the learning assignment.

2.2 The Dual task approach

On the dual task approach, Brünken et al., (2003) distinguish two different ways of use considering either a primary task approach or secondary task approach. Primary task approach supposes that a secondary task is added to a primary one to stimulate the memory load to perform the primary task. The variable of interest here is performance on primary task, which would decrease in reaction to a higher cognitive load. Memory and cognitive loads increased by a secondary task load would impact a decrease on the performance on the main task. Built on the same line of reasoning, the secondary task approach uses the second task to measure the memory load induced by the primary task on the secondary one. The variable of interest in this case is the performance on the secondary task (Brünken et al., 2003).

In the present study, we apply the primary task approach to measure the impact of a secondary task (Facebook usage) on the primary task (learning process). The working load grows up with a secondary task inversely stimulating the memory load on the main (primary) task. This fact induces a decrease of the student performance due to split attention on occupations, performing studying and surfing on the social website simultaneously. Focus is on keen interest of students to the networking site, intensity of its use and content drawn on during that time.

Further, we develop the research model to address the research objectives.

3 Research Model

3.1 Antecedents to Facebook usage and cognitive absorption

3.1.1 Personality traits

Personality has been brought to mind as for people's use of Facebook and its prediction to students' performance (Canales et al. 2009). Previous researches posit that personality has an important effect on students' results (Nofle and Robins, 2007, DeRaad and Schouwenburg, 1996; Digman and Takemoto-Chock, 1981). Distinct from a person's own intelligence, the personality influences students' academic performance (Nofle and Robins, 2007; Wagerman and Funder, 2007; Duckworth and Seligman, 2005). Several studies (Chamorro-Premuzic and Furnham, 2003; Rindermann and Neubauer, 2001; Blickle, 1996; Cacioppo et al., 1996; De Raad and Schouwenburg, 1996; Wolfe and Johnson, 1995; Goff and Ackerman, 1992) have been carried out with both college and universities students to analyse whether personality traits predict students' performance and achievements; controversial results have been concluded. While several earlier researches (Allik & Realo, 1997; Rothstein et al. 1994; Dollinger and Orf, 1991; Green et al. 1991; Mehta & Kumar, 1985) have concluded that personality doesn't significantly determine students' academic performance, more recent researches (Poropat, 2009; Tabak et al, 2009; Nofle and Robins, 2007; Conard, 2006; Chamorro-Premuzic and Furnham, 2003; Wolfe and Johnson, 1995) have argued that personality attributes are conclusive toward a better achievement in class. In fact, personality traits are associated to the ability of the person to do something, for instance to learn, and how the learner goes with the learning process. This stands for consistent evidence of the impact personality has on learners' results.

To this extent, where low agreeableness; high openness and high consciousness were initiated as predictors to achieve high scores; neuroticism, high agreeableness and extroversion pulled down the students' academic performance (Poropat, 2009; Tabak et al, 2009; Nofle and Robins, 2007; Conard, 2006; Chamorro-Premuzic and Furnham, 2003; Wolfe and Johnson, 1995). Indeed, Conard, (2006); Lounsbury et al. (2005) and Gray and Watson, (2002) studies have found that high level of **agreeableness** would favourably carry socialisation. People at this trait are straightforward, sympathetic and trustful (John and Srivastava 1999) which make easy for them to introduce themselves to new people they would meet on one website or the other. Their altruism, forgiving and tender-mindedness attitude facilitate their socialisation and they by far integrate social networks nowadays. Yet they might keep cautious attitude in terms of their interaction and content they provide on these websites.

Extroversion proves a quite similar attitude to agreeable people. However extroverts show more assertiveness, energy and excitement toward their surroundings. With their talkative character, outgoing stance and enthusiastic behaviour, they tend to be very sociable and adventurous than introverted but also agreeable people. Wiggins (1979), John and Srivastava (1999) and Judge et al. (2002) argue that extroverted people are dominant and effective in leading groups. Their agent character supports their success in this task. Extroverted people have self-confidence and are determinant and open which facilitate expressing their ideas, opinions and feelings more comfortably. Canales et al. (2009) identified that extroverted students have lower grades as a measure to their academic performance. Focus in the two studies is on how students are carrying out within their classes and difference in the conclusions can be explained by the level of attention given to the main task to perform. Extroverted students would be easily diverted from the main task of learning, being more attracted by an activity that absorbs their impulsiveness and excitement seeking. Moreover, Canales et al. (2009) study's outcome shows that extroverted (and neurotic) people are heavy users of Facebook suitably to their gregariousness and assertiveness. Thus, they would tend to be the owners of a big number of groups on this so-

cial network that lead discussions about many sorts of subjects that interest the audience of this community. Furthermore, as a deduction from Judge et al. (2007) and Canales et al. (2009) studies results, extroverted students tend to invest on time resources to spend more on social networking than studying. High mental effort is devoted to social activities such as group and instant discussions, video and pictures uploading and sharing which declines the mental effort allocated to the learning tasks.

Controversially, **conscientious** people as representation of directivity, self-discipline, organization and efficiency would better organize their involvement in more than one activity at once. Students with this specific character are deliberate and feel dutifulness, thus strive for achieving their objectives, for instance good academic results. Their sense of order, carefulness and competence make them methodical and scrupulous in their time management and the way they handle the tasks to perform. Students at this trait are more rigorous regarding their studies and feel responsible to accomplish expected results and can't tolerate to be careless. Being hardworking and persistent, conscientious students manage to achieve a high academic performance (Wolfe and Johnson, 1995, Judge and Ilies 2002, Bar-chard 2003; Conard 2006, Tabak et al. 2009)

Neuroticism as a negative nature would have a negative impact on performance (Barrick et al. 2001, Tabak et al. 2009). The specific personality character is expresses an unconstructive and unproductive temperament. People who regularly feel anxious, worried, not self-confident and shy tend to be introverted and participate only in activities that go with their personalities. They are either performing activities on their own, or unenthusiastically impact a group work. They either do not or rarely participate in social activities. With their continuous feeling of irritation and depression, they are more likely to do not choose to integrate a social community. However, with their character of vulnerability due to lack of self-confidence, they might tend to join social networking in answer to their impulsive, moody character and propensity to achieve content impression.

While previous studies have shown that neurotic people tend to over use internet (Hardie and Tee 2007, Hamburger and Ben-Artzi 2000), with the consideration of the interaction character within the social websites, we presume that neurotic people would either have passive presence on these websites.

Preliminary results from Canales et al. (2009) show that neurotic people use of Facebook positively correlates with the time spent on this social website, but this doesn't tell more about the extent of use and whether they have been interacting with others on the social website or not.

Curious people or people who are **open** to experience are imaginative, have large set of interests and like to explore new ideas. However, this doesn't necessary imply that they have an outgoing attitude or are adventurous. In fact, they act in terms of interests and curiosity about new ideas, with excitation and fantasy to added value they can get out of this openness to experience. Though, people with this personality trait are unconventional and stick to their values.

Referring to these definitions and outcomes from prior research, we hypothesize that:

H1a: Personality traits has an impact on Facebook usage and

H1b: Personality trait has an effect on cognitive absorption on Facebook.

3.1.2 Basic values

Largely discussed and used in political attitudes formation context, basic personal values refer to people's values and subjective beliefs (Schwartz, 1992). Schwartz (1992) categorizes personal basic val-

ues into four dimensions which illustrate what people assume in their lives: with openness to change people are sociable and outgoing. They are determined and easily adopt changes to achieve objectives they set. Self transcendent individuals make their individual perceived success out of help they would be able to provide to others and support they can bring for any one that needs assistance. Self enhanced people are people full of inner energy more often willing to be contagious to each individual in continuous social interaction with them. Yet, conservative individuals need rather security, so they stick to tradition and conformity to principles (Schwartz, 1992, 2006, 2010, 2014).

We hypothesize that:

H2a: Personal basic values have an impact on Facebook usage and

H2b: Personal basic values have an impact on cognitive absorption.

3.1.3 Social status

Social status refers to a person's social position considering its wealth, occupation and education level (Adler and Stewart, 2007). While Adler and Stewart (2007) identify two ways to measure social status, objective and subjective ways, both tools are complementary, so they do not present the individual's information redundantly. While objective social status is measured by a tier party recognition of the family and individual wealth, occupation and education level, subjective social status reports these indicators adding respective weight to what should be valued more among the three dimensions.

Adler and Stewart (2007) posit to use MacArthur ladder for individuals to position themselves in comparison to their respective population which is a self-reported ranking and measure. However, objective social status scale present quantified measures of one's financial status, job position and education level. In these conditions, social status would define individuals' access to internet and social networking platforms in terms of devices used, accessibility of the internet service 24hours a day and individual availability to go on social networking sites.

We presume these attributes to affect students' presence on the social networking platforms, its duration and cognitive immersion to determine level of control one has on its presence on the social networking site, how enjoyable this experience gets, the number and nature of activities performed and recognition of the experience associated to the time spent on the platform.

Thus, we hypothesize that:

H3a: High social status would increase Facebook usage.

H3b: High social status would increase the cognitive absorption on Facebook.

3.2 Facebook usage, cognitive absorption and students' achievements

Social Media (SM), social networks or social networking websites; expressions seem to have different meanings, but refer to a common concept, a precursor that conducted the worldwide into a change allowing individuals to communicate instantaneously via networks (Lariscy, et al. 2009). Previously, tools that are used for social exchange, such as: email, intranet, blogs, videoconferencing, podcast, video sharing, instant message, events, social networking, text messaging, photo sharing, wikis, virtual worlds, social bookmarking, gaming, micro-blogging (presence applications) and news aggregation have been also call SM (Erych et al., 2008). In this study, by SM we refer more specifically to social

networking websites and more specifically to Facebook. This system describes in itself the online practices that make use of technology and enable people to share content, opinions, experiences, insights, and technologies themselves. An illustration can be information shared via Twitter, videos shared on Youtube or Facebook or contacts and comments made on Facebook that point up the use of these new systems. Facebook as one of these social networks is today the top ranked in terms of use (Alexa Traffic Rank, 2010).

Karpinski and Duberstein's exploratory study (April 2009) posits that students spend from 3 to 4 hours a day on Internet and more than 5 hours per month surfing on Facebook. Statistics from the social media guide Mashable (October, 2009) show that people are spending about 5hours 39 minutes per month on Facebook. More recent study (November 2009) shows that Facebook surfers spend 3 full days per year on Facebook; which is equivalent to 6 hours per month. Statistics from the Facebook website itself (2010) points that 50% of their users that count over 400 millions, log on to the social website each day. Besides, debate wise from the International debate education association (IDEA) announces on February 2010 that facebookers spend 2.6 millions minutes (equivalent to about 43 334 hours) on the social network each day. This can be translated by an average time of 6 hours 30minutes per month per person. More recent statistics, (kissmetrics, 2015) show that average time spent on the social platform is 700min (11hours and 40mn per person per month). Wearesocial.com (2015) notes that average individual presence on Facebook is 4hours and 25min per day.

This is demonstrating a continuous increase of the usage of Facebook across the globe. Understanding how a computer, internet, (and more specifically Facebook use), can enhance or detract the students' achievements is the responsibility of both parents and teachers (Shields and Behrman, 2000) that need to understand this phenomenon. Furthermore, authorities are concerned and are wondering about the consequences to this extensive use for different groups of people. Besides, two recent studies (Karpinski and Duberstein, 2009 and Canales et al. 2009) have confirmed a negative impact of the time spent on this social network and their performance measured through their reported grade-point averages (GPAs). However, both studies have been run in US and focus was on the time spent on the social network, although Canales et al. (2009) spotlighted the role of the personality in determining the level of use of this social media.

As presented above, we presume that Facebook usage has a negative effect on students' achievements due to students' split attention on dual tasks. Students who are spending long hours a day surfing Facebook and exchanging with their friends online are taking time from their duties for the courses.

Furthermore, IT systems usage has been vastly discussed. This concept has had wide range of understandings and complex approaches (Venkatesh et al. 2008; Burton-Jones and Straub 2006; Jasperson et al. 2005; Straub et al. 1995). Technology use and Information system (IS) use has been the largest discussed approaches in the literature and business press. Reflections were mainly focused on information systems deployment and their success measurement through new technology use concept.

Venkatesh et al. (2008) highlight the importance of technology use itself but also the surrounding practices. Depending on the purpose of the system use measurement and the context of the system use itself, there are different approaches to represent the system use. Most common variables that apply to our context are: extend of student involvement in the use of this system, the duration of such use and the frequency.

In terms of cognitive absorption, Agarwal and Karahana (2000) have identified cognitive absorption as the construct that put together

Building upon this line of argument, we presume that:

H4a: Usage of Facebook has a negative impact on students' achievements, and

H4b: Cognitive absorption has a negative impact on students' achievements.

We hypothesize also that high cognitive absorption would increase Facebook usage. Thus we test the following hypothesis:

H5: High cognitive absorption would increase Facebook usage.

3.3 Moderating effect of polychronicity

Polychronicity or multitasking has been widely discussed in previous researches mainly focused on time management in job performance (Hecht and Allen 2005; Kaufman-Scarborough and Lindquist, 1999, Slocombe and Bluedorn, 1999; Lee, 1999). The concept has been largely introduced as a cultural characteristic that varies across the countries. It relates to the way people of one country or the other are used to manage their daily activities (Lee, 1999). While monochronicity is about performing one task at once and scheduling tasks to be run one after the other, polychronicity is essentially about pairing at least two tasks at once (Hecht and Allen 2005; Kaufman-Scarborough and Lindquist, 1999; Lee, 1999). Therefore, we presume polychronicity varies cross countries and set the following hypotheses that presume the moderating effect of polychronicity on the impact of Facebook usage on the students' achievement as well as its moderating effect on the impact of cognitive absorption on students' achievement. We check polychronicity capacity across three countries sample. Then, we test whether a student with higher polychronicity ability would be less vulnerable to the negative impact of Facebook usage on his academic performance.

Thus, we hypothesize that:

H6a: Polychronicity level has a significant moderating effect on Facebook usage impact on students' achievement.

H6b: Polychronicity has a significant moderating effect on cognitive absorption impact on students' achievement.

The research model below (Figure 1) depicts the interdependencies defined above. Further, we describe the empirical study to be run with students who use Facebook.

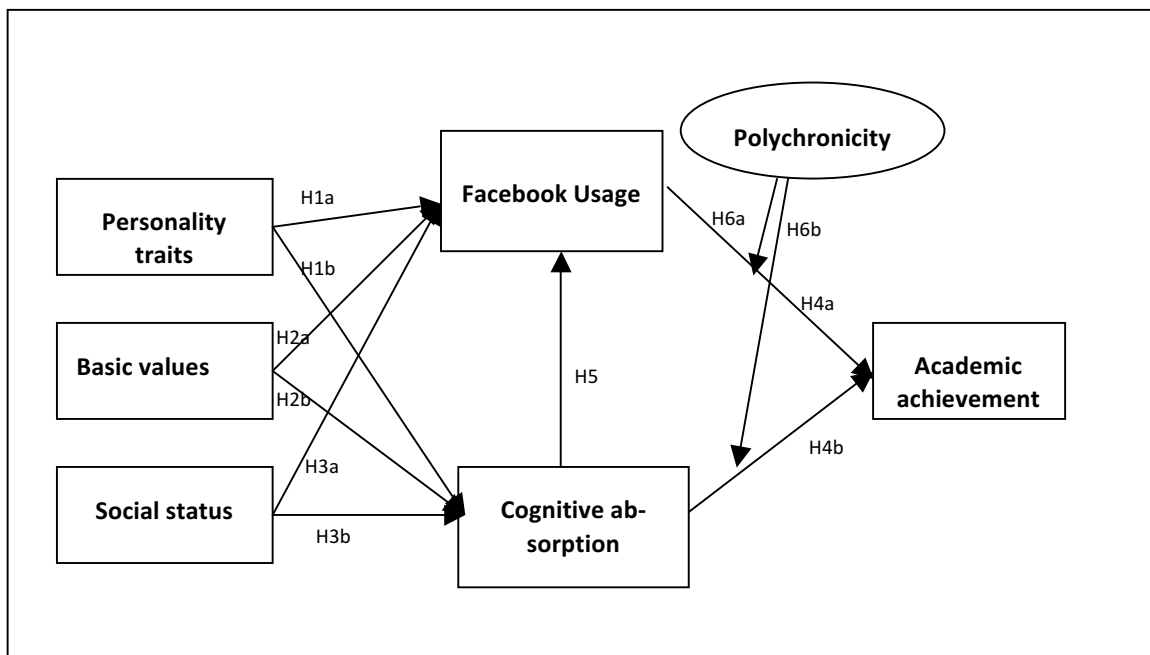


Figure 1. Research Model

4 Methodology

4.1 Data collection

A survey is being designed and will be sent to undergraduate business students from three schools of Management in three countries, France, Kingdom of Saudi Arabia and Tunisia.

4.2 Measures

Measures used in the survey are adapted from confirmed studies in social and educational sciences where prior researches have shown reliability and validity in their respective contexts.

Yet, some changes occurred on the measures' wording in order to adapt the scales to the context of our study. Changes have mainly concerned "Facebook usage" construct. Measures have been used from "the Information System usage" at the Information System (IS) discipline.

Examples from the measures used are presented in the table 1.

| Constructs | Measures | Reference |
|---------------------------------|--|--|
| Facebook usage | Time spent Frequency Extent (activities) | Burton-Jones and Straub (2006) |
| Cognitive absorption on the web | Temporal Dissociation Time appears to go by very quickly when I am using the Web, Focused Immersion While using the Web I am able to block out most other distractions. Heightened Enjoyment I have fun interacting with the Web. Control When using the Web I feel in control Curiosity Using the Web excites my curiosity | Agarwal and Karahanna (2000) |
| Personality traits | Extroversion I like to have a lot of people around me. Agreeableness I try to be courteous to everyone I meet. Conscientiousness I strive for excellence in everything I do. Neuroticism I often get angry at the way people treat me. Openness I have a lot of intellectual curiosity. | Costa and McCrae (1992) |
| Polychronicity | I do not like to do several activities at the same time I should not try to do many things at once | Kaufmann- Scarborough and Lindquist (1999) |

Table 1: Measures examples

4.3 Data Analysis

Collected data will be analyzed using structural equation modeling statistical method and SmartPLS package to test measurement and structural models. Convergent and discriminant validities will be measured for each construct and evaluate the significance level of hypothesized causal effects.

4.4 Expected contributions

The research study is expected to make contribution at two main levels:

- Develop a research model that account of combined personal characters that explain Facebook usage and online cognitive absorption.
- We will assess satisfaction with life variance explained by the research model. We will also analyze condensed effect of personality, basic value and social status and quantify explained variance of Facebook usage and cognitive absorption by respective individuals' profiles.

5 Conclusion

Use of the social networks has been expanding very widely with the use of the social media such as MySpace, Youtube, Twitter or Facebook. Preliminary studies have analyzed the impact of personality traits on the use of these networks and its preliminary impact on the students' achievement. However,

little consideration has been given to the cognitive load induced by the pair run of both homework and interaction on the social websites. Lower reflection toward the impact of the mental effort the students devote to their studies was also stated. Taking into account these unexplored area to explain the real impact of the SM use on students' results, we presumed that simultaneous use of Facebook and homework preparation stimulates a high extraneous load that negatively impacts the working memory load. While a high cognitive load is experienced lower grades result.

The designed comparative study will allow comparing academic achievement for students with similar personal characters although in different cultural background settings, compare results from three countries two by two and assess to which extent Facebook usage and cognitive absorption are interdependent and would affect students' academic achievement while controlling the final results through the polychronicity capability level.

References

- Adler, N. and Stewart, J. (2007) 'The MacArthur Scale of Subjective social status'. Accessible at: <http://www.macses.ucsf.edu/research/psychosocial/subjective.php#measurement>. (Visited on 07/12/2015).
- Agarwal, R. and Karahanna, E. (2000) Time Flies When You're Having Fun: Cognitive Absorption and Beliefs About Information Technology Usage, *MIS Quarterly*, 4(1), 665-694
- Alexa Traffic Rank, June 2010. Retrieved from: <http://www.alexa.com/siteinfo/facebook.com>. Visited on 07/07/2010).
- Barrick, M.R. and Mount, M.K. (1991). The Big Five Personality Dimensions and Job Performance: A Meta Analysis. *Personnel Psychology*, March, 44(1), 1–26. doi: 10.1111/j.1744-6570.1991.tb00688.
- Brünken, R.; Plass, J.L. and Leutner, D. (2003), "Direct Measurement of Cognitive Load in Multimedia Learning", *Educational Psychologist*, 38(1), 53–61
- Canales, C., Wilbanks, B. Yeoman, A. (2009). Does Facebook have negative impacts on academic performance? Facebook usage in relation to personality and academic performance. URL: <http://www.covenant.edu/news/09.28.09>. Visited on 06/06/2011.
- Eyrich, N. Padman, M.L. and Sweetser, K.D. (2008), "PR practitioners' use of social media tools and communication technology", *Public Relations Review*, 34, 412–414,
- Hardie, E. and Tee, M. Y. (2007). Excessive Internet Use: The Role of Personality, Loneliness and Social Support Networks in Internet Addiction. *Society*, 5(1), 34-47. URL <http://hdl.handle.net/1959.3/5548>. Visited on: March 02, 2011.
- Karpinski, A.C. and Duberstein, A. (2009), A description of Facebook use and academic performance among undergraduate and graduate students. American Educational Research Association (AERA) Conference, San Diego, 13-17 April.
- Lariscy, R.W., Avery, E.J., Sweetser, K.D and Howes, P. (2009), "Research in brief: An examination of the role of online social media in journalists' source mix", *Public Relations Review* 35, 314–316.
- Paas, F. (1992) "Training Strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A cognitive Load Approach", *Journal of Educational Psychology*, 84(4), 429-434.
- Paas, F. and Van Merriënboer, J.J.G. (1994), "Instructional control of Cognitive load", *Educational Psychology Review*, 6(4), 351-371.
- Paas, F.; Renkl, A. and Sweller, J (2003a), "Cognitive Load Theory and Instructional Design: Recent Developments", *Educational Psychologist*, 38(1), pp.1-4.
- Paas, F; Tuovinen, J.E.; Tabbers, H. and Van Gerven, P.W.M. (2003b), "Cognitive Load Measurement as a Means to Advance Cognitive Load Theory", *Educational Psychologist*, 38(1), 63–71
- Schwartz et al. (2014) "Basic Personal Values Underlie and Give Coherence to Political Values: A Cross National Study in 15 Countries" *Political Behavior*, 36(4), 899-930, December.

- Schwartz S.H., Caprara G.V and Vecchione M. 2010. 'Basic Personal Values, Core Political Values, and Voting: A Longitudinal Analysis'. *Political Psychology*, 31(3), 421-452.
- Schwartz, S. H. 1992. "Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries". *Advances in experimental social psychology*, 25, 1–65.
- Schwartz, S. H. 2006. 'Basic human values: Theory, measurement, and applications'. *Revue Française de Sociologie*, 47, 249–288.
- Tabak, F., Nguyen, N., Basuray, T. and Darrow, W. (2009). Personality and individual differences exploring the impact of personality on performance: How time-on-task moderates the mediation by self-efficacy. *Personality and Individual Differences*, 47, 823–828.
- Venkatesh, V., Brown, S.A., Maruping, L.M and Bala, H (2008), "Predicting Different Conceptualizations of System Use: The Computing Roles of Behavioral Intention, Facilitating Conditions, And Behavioral Expectation", *MIS Quarterly*, 32(3), 483-502, September.
- Venkatesh, V.; Morris, M.G.; Davis, G.B and Davis, F.D (2003), "User Acceptance of Information Technology: Toward a Unified View, *MIS Quarterly*, 27(3), 425-478, September.
- Wearesocial.com (2015), Digital, Social & Mobile Worldwide in 2015, URL: <http://wearesocial.net/blog/2015/01/digital-social-mobile-worldwide-2015/>. (Visited on 06/20/2015).

FACTORS AFFECTING RECOMMENDATIONS' ACCEPTANCE IN OFF-LINE ENVIRONMENT

Complete Research

Gkika, Sofia, ELTRUN: The e-Business Research Center, Athens University of Economics and Business, Greece, gkikas@aueb.gr

Skiada, Marianna, ELTRUN: The e-Business Research Center, Athens University of Economics and Business, Greece, mskiada@aueb.gr

Lekakos, George, ELTRUN: The e-Business Research Center, Athens University of Economics and Business, Greece, glekakos@aueb.gr

Abstract

Recommender Systems are traditionally been used in online environments. The present study focuses on factors affecting consumers' acceptance in recommendation systems that will be used in traditional stores. For this purpose, we have developed a garment recommendation system applying and evaluating the aforementioned factors through an experiment. The main insights of the study indicate that consumers' personality affect consumers' preferences and garments' style.

Additionally, when a consumer has high motivation to purchase a particular garment (s)he does not waste his/her time for looking around other types of garments. This means that if a garment provider wants to increase cross selling then he has to recommend products that are close to that event such as accessories.

Keywords: *recommender systems, persuasion, acceptance.*

1 Introduction

Recommender Systems are typically applied in online environments and are characterized as one of the most successful software tools since they meet the needs of both product/service providers and customers. The main inputs of a Recommender System in an on-line store are users' id, his/her rating upon products (s)he has seen and purchase history. However in a physical environment data collection is more complex since the data that should be collected for the provision of a recommendation through a kiosk, is a combination of data from both off-line (e.g. which products (s)he holds) and on-line environment (e.g. user's id, history purchase). The present study focuses in that point which is it's main distinction from other studies, the design of a Recommender System for a physical (off-line) store. Given that factors that affect consumers' behaviour when they are shopping on a physical store (such as time pressure, crowding, etc) are still to be investigated, it would have been interesting to study in that direction. At the first phase, our goal is to determine the factors affecting recommendations' acceptance and consecutively to implement them on a physical store.

The context of this study is physical garment store. The design of a recommendations' service that will suggest garments (or other accessories) to consumers according to their preferences is one of the innovative consumer services that are willing be designed in the fashion domain. Recommending garments and relevant accessories to consumers gives product providers the opportunity to increase sales as well as to increase satisfaction loyalty of their customers. However, in order to secure the success of the recommendation service, it is crucial to identify the factors that affect the consumer's decision making towards using the recommendations services in the in-store environment.

Thus, we will focus on a Recommender System which recommends garments, examining it – at this stage of the project - from the interface perspective. More specifically, we will focus on investigating the impact of three important factors on the acceptance and use of the recommendation service is consumer's personality. For this purpose, we have developed a garment recommendation system applying and evaluating the aforementioned factors through an experiment, as described in the next sections.

The paper is organized in five sections. In Section 2 is described the role of persuasion and the factors affecting recommendations' acceptance. Section 3 presents in detail our study which is divided into two phases the garments' style classification and the experiment. In Section 4 are presented the experimental results are discussed, while in Section 5 the main conclusions of this research are demonstrated and what our future work involves.

2 Literature Review

2.1 Factors affecting recommendations' acceptance

The main research in Recommender Systems has traditionally been focused on their algorithmic perspective. The implicit assumption that the more accurate recommendations are provided, the more successful the RS is, has recently been challenged since other factors have been identified as antecedents of a RS success. Some of these factors are the presentation of a recommendation (Nanou et al. 2010), the set size of recommendations (Ho and Tam, 2005), recommendation's novelty, serendipity (i.e. Pu and Chen, 2011, Cremonesi et al., 2010) as well as user's personality (Gkika and Lekakos, 2014).

Clothing style preferences differentiate among humans. Body type, body image, age as well as fit preferences are some of the factors that affect someone's clothing style (Chattaraman and Rudd, 2006; Chattaraman et al., 2013). In order someone to feel comfortable with his/her clothing, usually prefers to purchase garments from a specific clothing style.

H1: Consumer's style affects garment preferences

H2: Consumer's purchase intention is affected by garment's clothing style

A human's personality is defined as 'a dynamic organisation, inside the person, of psychophysical systems that create the persons' characteristic patterns of behaviour, thoughts and feelings' (Allport, 1961, p. 11). Studies indicate that personality influences the way people make their decisions (Nunes et al., 2012). Given that the type of services we examine are trying to form, change or reinforce people's opinions and decisions, they are trying to persuade them in other words, it would have been useful to examine how people with different personality are getting persuaded. According to Moddy et al. (1977) there is correlation between three out of five personality factors and clothing style preference. Moreover, studies have shown that dress express personality, nevertheless clothing choice has also been viewed as overt behavior, which is being influenced by an individual's personality profile (Gurel et al., 1972). Thus, the third hypothesis of the experiment is:

H3: Consumer's personality affects her intention to purchase garments

The Elaboration Likelihood Model (ELM) indicates that individuals with low motivation or ability to process the information provided with a recommendation could eventually get persuaded to select/use the item if appropriate peripheral cues enrich the recommended item. According to the Fogg (2009) in order for someone to get persuaded to perform certain behaviour (s)he should have motivation to act, the ability to do it and get prompted by an activating trigger. If any of those three components is not present or their mix is not appropriately balanced then the individual will not perform the target behaviour. In case someone lacks motivation then the triggers should be in the form of 'motivational elements' so as to grab customer's attention and consecutively make them see and process the provided recommendations.

H4: Consumer's motivation to purchase affects the amount of products (s)he purchase.

In order to evaluate the above hypotheses an experiment was conducted as described in the next section.

3 Research Design and Methodology

3.1 Garments' style classification

In order to investigate whether people with the different personality have different clothing style (H1) preferences, garments had first to be categorized regarding to their style so as to examine which clothing style(s) is/are preferred from each personality group. In order to experiment with a wide variety of garments, different style garments have been used. We adopted Shen's et al. (2007) methodology and taxonomy of clothing style, which consists of 6 categories (formal, luxurious, trendy, funky, casual and sporty). Shen et al. (2007) indicate that a garment may belong to more than one type of clothing, for instance a black midi dress may be characterized both formal and luxurious.

For the classification of garments in our experiment, 14 fashion experts evaluated 35 garments from the last Diffusione Tessile's collection, an Italian garment provider. For each garment the experts were asked to evaluate each garment's compliance with the six style groups and providing a rating from 0-10 (0 not matched to a category – 10 perfectly matched to a category). Finally, 24 out of 35 garments were selected and used in our experiment, those with the highest matching averages.

3.2 Experiment Design

In order to investigate the above hypotheses a between-groups experiment was conducted for two different scenarios of the customer experience use case. The first scenario simulates a low intention/ low motivation shopping behavior, where consumers visit the store without a clear intention to buy something. This "Shopping therapy" scenario differentiates from the second scenario, which simulates a high intention/high motivation scenario where consumers visit the store in order to perform a planned purchase (e.g. a garment for an important event). The distinction of the two scenarios was necessary since (according to the Elaboration Likelihood Model) users with different level of motivation levels respond differently to motivational triggers with respect to the acceptance of recommendation. The first scenario asked the participant to imagine that it is Saturday morning and she is going for "shopping therapy" while at the second scenario she is supposed going shopping in order to purchase a garment for an important event to her (e.g. a special dinner, a date, an appointment) which would take place that day. Consecutively, in both scenarios, experiment participants were asked to imagine that they enter a store in which a number of recommendations (24 in total) were provided to them through a touch screen monitor.





At the first step of the experiment, participants were asked to fill-in a questionnaire with personality related questions, utilized to classify them in personality groups in order to examine potential relationship between personality and garment preferences (style groups) as well as the acceptance of recommendations (H3/H4).

There is a variety of tools and methodologies in order to measure a person's personality but the most widely accepted and used is that of the Five-Factor Model of personality, or the Big Five (John, 2008). According to Five-Factor Model, each person is characterized by five personality traits (extraversion, neuroticism, agreeableness, conscientiousness and openness to experience) to lower or higher degree. The questionnaire that used for personality is Big Five Inventory (John et al., 1991; 2008) which consists of 44 questions. Moreover, there were questions about users' motivation to purchase (Tam and Ho, 2005). The particular questions were used so as to investigate whether people with a different personality profile behave in a different way in case of having both high and low motivation to purchase.

At the second step of the experiment, six blogs of garments were presented to the participants (i.e. six web pages with four garments each) (Figure 1 and Figure 2). The participants were asked through a questionnaire if they like the recommended garment and match their style (by providing ratings in a 1 to 5 scale) as well as whether they believe the garment is novel or a serendipitous recommendation for them and finally their intention to purchase it or not. The above questions are concerning novelty and serendipity used in previous studies (Celma et al., 2008; Adamopoulos & Tuzhilin, 2013). The goal of the aforementioned questions was to examine whether a person with a particular personality profile considers a recommendation (a garment with a specific clothing style) novel or serendipitous as well as her intention to purchase it.



It is Saturday morning and you are going for "shopping therapy". You enter a store with women garments and recommendations are provided to you through a touch screen monitor.
 Assume that you have sufficient amount of money for your shopping.
 (Answer the following scenario-based questions)

| | | | |
|--|--|--|--|
|  |  |  |  |
| I like this garment (1(not at all)-5(very much)) 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> | I like this garment (1(not at all)-5(very much)) 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> | I like this garment (1(not at all)-5(very much)) 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> | I like this garment (1(not at all)-5(very much)) 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> |
| This garment matches my style (1(not at all)-5(very much)) 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> | This garment matches my style (1(not at all)-5(very much)) 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> | This garment matches my style (1(not at all)-5(very much)) 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> | This garment matches my style (1(not at all)-5(very much)) 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> |
| I would buy this garment (under this scenario) Yes <input type="radio"/> No <input type="radio"/> | I would buy this garment (under this scenario) Yes <input type="radio"/> No <input type="radio"/> | I would buy this garment (under this scenario) Yes <input type="radio"/> No <input type="radio"/> | I would buy this garment (under this scenario) Yes <input type="radio"/> No <input type="radio"/> |
| I was surprised by this recommendation Yes <input type="radio"/> No <input type="radio"/> | I was surprised by this recommendation Yes <input type="radio"/> No <input type="radio"/> | I was surprised by this recommendation Yes <input type="radio"/> No <input type="radio"/> | I was surprised by this recommendation Yes <input type="radio"/> No <input type="radio"/> |
| I have never seen this garment before Yes <input type="radio"/> No <input type="radio"/> | I have never seen this garment before Yes <input type="radio"/> No <input type="radio"/> | I have never seen this garment before Yes <input type="radio"/> No <input type="radio"/> | I have never seen this garment before Yes <input type="radio"/> No <input type="radio"/> |
| See next recommendations (1/6) | | | |

Copyright © 2015

Figure 1: Second step of the experiment ('Shopping Therapy' scenario)



Figure 2: Second step of the experiment ('Event' scenario)

At the third and last step of the experiment a final questionnaire was provided which contained questions a few demographic questions.

3.2.1 Sample

The experiment participants were invited through posts in University's Facebook groups (e.g. undergraduate, postgraduate and PhD students). The invitation message was asking recipients to participate in a research in which they would be asked to rate recommendations provided by an application as well as to fill in a psychographic questionnaire in the domain of fashion. The link to access the system was provided and a clear suggestion concerning the anonymity of their participation was included in the message.

The properly completed surveys were 38 in the 'Shopping Therapy' scenario and 33 in the 'Event' scenario, in a total of 71 participants. All users were females while the 46% of the sample were aged between 18 and 24 years old, the 52% were between 25 and 34 years old and the 2% at the age of 35-44 years old.

3.2.2 Survey Results

The first step of our analysis was to investigate the effect of consumer's personal clothing style on her garment preferences (H1). For both scenarios we measured the participants' average preference rating for the garments on each one of the six clothing styles and then an ANOVA analysis was conducted in order to measure potential differences among the different styles. ANOVA results suggested that in both scenarios, there are significant differences ($p < .001$) among the average rating of garments in each clothing style category (Table 1). Thus, H1 is confirmed.

| Scenario | Sign | Mean Square |
|------------------|------|-------------|
| Shopping therapy | .000 | 4.640 |
| Event | .000 | 3.294 |

Table 1: ANOVA results for H1

Consecutively, we compared in both scenarios, the ratio of the amount of garments (s)he declared she would purchase to the amount of garments she declared that match her style (Case 1). Then, we compared it to the ratio of the amount of garments (s)he declared she would purchase to the amount of garments that does not match her style preferences (Case 2). In this way we investigated whether the consumer's intention to purchase was affected by the garment's clothing style (H2). A paired t-test analysis was performed and the results indicated that the difference of means between the above elements is positive (Table 2), which denotes that a garment's style affects the consumer's intention to purchase it.

| Scenario | Case 1 – Case 2 | | |
|------------------|-----------------|---------------|-------|
| | Mean | St. Deviation | Sign. |
| Shopping therapy | 719 | .278 | .000 |
| Event | .649 | .281 | .000 |

Table 2: T-test Results for H2

In order to investigate H3, we examined which combinations of personality traits tend to score highest ratings on a diverse group of garments' style. Thus, the fsQCA methodology was performed for both scenarios. The five personality traits were used as possible conditions that influence the acceptance of garments (in the form of liking the garments). As a first step, the prescriptions of fsQCA required calibration of the cases into membership sets through fsQCA 2.0 soft-

ware. The function demands as input three threshold points; a full-membership value, a non-membership value and a cutoff point. Because the dataset consists of subjective cases, we used cluster analysis following the k-means algorithm (k=3) to calculate the three membership sets. More specifically, high values are correlated with the full-membership set, medium values are correlated with the crossover point set and finally low values are correlated with the non-membership set.

The results of fsQCA indicate 3-7 alternative solutions comprising of alternative combinations of the personality traits that lead to high garment preferences. Black circles indicate the required presence of a personality trait in a solution. White circles indicate the required absence of a personality trait from the solution. Blank cells indicate that in that particular solution, the presence or absence of that personality trait is indifferent. Each solution is accompanied by two additional measurements of fitness, which express the 'predictive power' of each solution, namely the consistency and coverage indexes. Consistency presents how consistent is the empirical evidence with the outcome which is investigated while coverage estimates the proportion of cases that address the outcome which is under investigation.

Table 3 illustrates the results of fsQCA for the Shopping Therapy Scenario. The analysis identified seven solutions leading to high intention to purchase garments. All solutions require the absence of a personality trait by individuals. For example, the first solution proposes that individuals who are not agreeable but they are conscious and open to new experiences are likely to have high intention to purchase garments.

| | Solutions leading to high intention to purchase Garments | | | | | | |
|------------------------------|--|-------|-------|-------|-------|-------|-------|
| Personality Traits | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Extraversion | | | | ○ | ○ | ● | ● |
| Agreeableness | ● | ● | | | ○ | ● | ○ |
| Conscientiousness | | ○ | ○ | ○ | ● | ● | ● |
| Openness | ○ | ○ | ○ | ○ | ● | ● | ○ |
| Neuroticism | ○ | | ○ | | | ● | ● |
| Consistency | 0.7 | 0.682 | 0.648 | 0.651 | 0.892 | 0.956 | 0.875 |
| Coverage | 0.43 | 0.421 | 0.654 | 0.666 | 0.189 | 0.155 | 0.183 |
| Overall solution consistency | 0.635 | | | | | | |
| Overall solution coverage | 0.812 | | | | | | |

Table 3: fsQCA results for the paths leading to intention to purchase garments (Shopping Therapy Scenario)

The next Table 4 presents the different paths, consisting of combinations of personality traits, which lead to high acceptance of garments in the case of Event Scenario.

| Personality Traits | Solutions leading to high intention to purchase Garments | | |
|------------------------------|--|-------|-------|
| | 1 | 2 | 3 |
| Extraversion | | ● | |
| Agreeableness | | | ○ |
| Conscientiousness | | | ○ |
| Openness | ● | | |
| Neuroticism | | | |
| Consistency | 0.887 | 0.853 | 0.671 |
| Coverage | 0.4 | 0.49 | 0.746 |
| Overall solution consistency | 0.790 | | |
| Overall solution coverage | 0.680 | | |

Table 4: fsQCA results for the paths leading to high intention to purchase garments (Event Scenario)

In order to examine whether the consumer's motivation to purchase affects the amount of products (s)he purchase (H4), we measured and compared, for both scenarios (Shopping Therapy Scenario – Event Scenario), the average means of the garments that the participants had declared their intention to purchase (Table 5). The paired t-test results indicated that there are significant differences ($p < .05$) for the aforementioned metric between the scenarios. Hence, we accept H4.

Shopping Therapy Scenario – Event Scenario

| Mean | St. Deviation | Sign |
|-------|---------------|------|
| 1.667 | 4.105 | .035 |

Table 5: T-test Results for H4

4. Main Findings and Future Work

The present research emphasizes the impact of consumer's personality on the acceptance and use of recommender systems. First, we empirically validated that a consumer's clothing style affects which garments (s)he is going to purchase. His/Her purchase preferences are close to his/her personal clothing style. Thus, when a recommender system provides products that are related to someone's style (such as garments or shoes) then the recommendations should be close to his/her personal preferences on the grounds that the more recommendations close to his/her style, the more purchases (s)he will intend to do. In fashion's domain terms, garments' grouping according to user's style is very important.

Moreover, the effect of human's personality on his/her garment preferences is validated in the present experiment. Due to the fact that our sample was quite limited we could have only an indication that people with different personality profile have different garment preferences. In the fashion domain, the consumer's personality profile can be captured by the recommender system through the customer's loyalty card. The customer may scan his/her loyalty card on touch-screen monitors and then personalized services and recommendations can be provided to the customers.

Recommendations won't be provided through only their purchase history (like traditional Recommender Systems), since personal clothing style and personality profile should also be taken into consideration for an effective recommender system.

Customers' sometimes go shopping for fun, they do not want to purchase something (low motivation to purchase) while some others go for planned purchases (high motivation to purchase) for a wedding or a special event. The results indicate that when (s)he has high motivation to purchase a particular garment (s)he does not waste his/her time for other types of garments. This means that if a product provider wants to increase cross selling then he has to recommend products that are close to that event such as accessories. In our future research we intend to concentrate on the impact of the types of recommendations (serendipitous, novel, diverse) on consumers' acceptance.

Acknowledgments

The present research is funded by the Project " SENSOR -ENABLED REAL WORD AWARENESS FOR MANAGEMENT INFORMATION SYSTEMS" code EP-2103-01. The Project co-funded by the European Commission within the Seventh Framework Pro-gramme (2007-2013).

References

- Avison, D. E. and G. Fitzgerald (1995). *Information Systems Development. Methodologies, Techniques and Tools*. 2nd Edition. London: McGraw-Hill.
- Adamopoulos, P., Tuzhilin, A.(2015) On unexpectedness in recommender systems: Or how to expect the unexpected. ACM Transactions on Intelligent Systems and Technology (TIST) - Special Sections on Diversity and Discovery in Recommender Systems, Online Advertising and Regular Papers, 5(4).
- Allport, G. W. (1961) Pattern and growth in personality. New York: Holt, Rinehart & Winston.
- Celma O., & Lamere, P. (2007). Music Recommendation Tutorial. Presented at the 8th International Conference on Music Information Retrieval, Vienna, Austria.
- Chattaraman,V., Rudd, (2006) Preferences for Aesthetic Attributes in Clohing as a Function of Body Image, Body Cathexis and Body Size. Clothing and Textiles Research Journal, 24(1).
- Chattaraman, V., Simmons, K.P., Ulrich, P.V.: (2013)Age, Body Size, Body Image, and Fit Prefer-

- ences of Male Consumers. *Clothing and Textiles Research Journal* 31(4), pp. 291-305.
- Cremonesi P., Garzotto F., Turrin, R.(2012) Investigating the Persuasion Potential of Recommender Systems from a Quality Perspective: An Empirical Study. *ACM Trans. Interact. Intell. Syst.*, 2(2), 11:1–11:41.
- Fogg, B. J. (2009) A behavior model for persuasive design. *Proceeding of Persuasive '09 Proceedings of the 4th International Conference on Persuasive Technology*.
- Gkika, S., Lekakos, G. (2014) The persuasive role of Explanations in Recommender Systems. *Proceedings of the Second International Workshop on Behavior Change Support Systems (BCSS2014)*.
- Gurel, L. M., Wilbur, J.C., Gurel, C. (1972) Personality correlates of adolescent clothing styles. *Journal of Home Economics*, 64(3), pp. 42-47.
- Tam, KY., Ho, SY. (2005) Web personalization as a persuasion strategy: An elaboration likelihood model perspective. *Information Systems Research*, pp. 271 - 291.
- John, O. P., Naumann, L. P., Soto, C. J. (2008) Paradigm shift to the integrative big-five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 287–310). NewYork: Guilford Press.
- Matsumoto D., Kouznetsova, N., Ray R., Ratzlaff C., Biehl M., & Raroque, J. (1999). Psychological culture, physical health, and subjective well-being. *Journal of Gender, Culture, and Health*,4, 1-18.
- Moody, W., Sinha, P. (2010) An Exploratory study: Relationships between Trying on Clothing, Mood, Emotion, Personality and Clothing Preference. *Journal of Fashion Marketing and Management*, 14 (1). pp. 161-179. ISSN 1361-2-26.
- Nanou, T., Lekakos, G., Fouskas, K.(2010) The effects of recommendations' presentation on persuasion and satisfaction in a movie recommender system). *Multimedia systems*, 16(4-5), 219–230.
- Nunes, M. A. S. N., Hu, R.(2012) Personality-based Recommender Systems: An Overview. *In: Proc. of the 6th ACM Conference on Recommender Systems (RecSys'12)*, pp. 5-6.
- Park, K. (1997) Individuality on fashion diffusion: differentiation versus independence. *Journal of Fashion Marketing and Management: An International Journal*, 2(4), pp. 352-360.
- Pu, P., Chen, L., Hu, R. (2012) Evaluating recommender systems from the user's perspective: survey of the state of the art. *User Modeling and User-Adapted Interaction*, 22(4-5), 317–355.
- Schrank, HL., Gilmore, DL. (1973) Correlates of Fashion Leadership. *The Sociological Quarterly*, 14(4), pp. 534-543.
- Shen, E., Lieberman, H., Lam, F. (2007) What Am I Gonna Wear?: Scenario-oriented Recommendation. In *Proceedings of the 12th International Conference on Intelligent User Interfaces*,

pp. 365-368.

THE VALUE OF IT: EXPLAINING THE STRATEGIC ROLE OF INFORMATION SYSTEMS FOR FAST GROWING SMES

Complete Research

Obwegeser, Nikolaus, Aarhus University, Aarhus, DK, nikolaus@mgmt.au.dk

De Araújo Burcharth, Ana Luiza, Fundação Dom Cabral, Belo Horizonte, BR,
ana.burcharth@fdc.org.br

Carugati, Andrea, Aarhus University, Aarhus, DK, andrea@mgmt.au.dk

Abstract

This article investigates fast growing and successful small and medium enterprises (SMEs). We focus particularly on the role and relevance of information systems (IS) skills and knowledge for their success. Although many theories account for the relationship between IS and organizational performance, they target mostly the context of large corporations, thus overlooking the special conditions that fast growing SMEs face. We specifically investigate fast growing SMEs for the following reasons: these companies are subject to constant organizational change due to their continuous growth and thus rely critically on having the right IS setup and strategy, while at the same time they are often too small and specialized for out-of-the-box enterprise IS solutions. Based on an interpretive multiple-case study design, we present three generic, archetypal IS strategies, ranging from full identification with and capitalization on IS to complete cost-based outsourcing of all IS related services. We analyze the inherent chances and risks associated with each of these strategies and present anecdotal evidence from the case companies to illustrate our findings.

Keywords: SMEs, E-skills, Strategic alignment

1 INTRODUCTION

The influence of information systems (IS) on the success and performance of companies has been widely recognized, and is an active field of research (Chan, Huff, Barclay, & Copeland, 1997; Chen, Worth, Preston, & Teubner, 2010; Melville, Kraemer, & Gurbaxani, 2004). The concept of IS alignment has been established in both research and practice and describes the need for a combined approach that integrates both business and IS goals to form an overall strategy (Henderson & Venkatraman, 1993).

In this paper, we present an in-depth analysis of how an organization's understanding and valuation of IS relates to its success. We decided to specifically investigate fast growing SMEs for the following reasons: these companies are subject to constant organizational change due to their continuous growth and thus rely critically on having the right IS setup and strategy. At the same time they are often too small and/or specialized for - or simply cannot afford - out-of-the-box enterprise IS solutions and "have rather limited free time to work on strategic issues, limited IT knowledge, and limited resources to spend" (Bernaert, Poels, Snoeck, & De Backer, 2014). The European Commission defines SMEs as "enterprises which employ fewer than 250 persons and which have an annual turnover not exceeding

50 million euro, and/or an annual balance sheet total not exceeding 43 million euro” (EU, 2005). SMEs are a very relevant research setting because they are considered to be the “backbone” of the economy, constituting up to 99% of the whole population of firms,, contributing 2/3 of private sector jobs and being responsible for more than 50% of overall value-creation within the EU (Muller, Gagliardi, Caliandro, Bohn, & Klitou, 2014). Nevertheless, researchers found that there is a gap of scientific literature investigating the otherwise well-known concepts of IS strategy and Enterprise Architecture in the context of SMEs (Bernaert et al., 2014; Margi Levy, Powell, & Yetton, 2002). This research project therefore helps to understand the relevance of IS skills and knowledge on the success of fast growing organizations.

Based on an interpretive multiple-case study design, we present and analyze the findings of ten interviews. Using inductive concept construction we are able to abstract three generic, archetypical IS strategies from our case data, ranging from full identification with and capitalization on IS (IT companies) to complete cost-based outsourcing of all IS related services. We analyze the inherent chances and risks associated with each of these strategies and present anecdotal evidence from the case companies to illustrate our findings.

The remainder of this article is structured as follows: We first provide an overview of the concept of IS strategy and the particular challenges of SMEs facing continuous growth. In the next section we describe in detail our research design and methodology, followed by a presentation of the results from the multiple case studies, our empirical observations and a discussion of our findings. Finally, we present concluding remarks that sum up our research and point to implications for practitioners.

2 STATE-OF-THE-ART

Information systems are broadly conceptualized as the combination of multiple factors that encompasses technology, people, processes and networks. They are investigated at different levels, including development, implementation, provision and operation (Davis, 2000). IS research is often embedded in organizational settings and therefore closely connected to organizational issues. A large number of practitioners is aware of the influence that IS has on the success of organizations (Worthen, 2007). Especially in times of difficult economic situations, decision makers are interested in understanding the value that IS can provide for their companies (Luftman & Ben-Zvi, 2010).

In an exhaustive review of the IS strategy literature (Chen et al., 2010, p. 237) define IS strategy as “the organizational perspective on the investment in, deployment, use, and management of information systems”. They stress that it is due to this broad conceptualization that IS strategy can be aligned with and provide valuable benefits for an organization’s business strategy. Moreover, they find that the scientific literature presents itself using three different understandings of IS strategy: IS strategy to support the business strategy, IS strategy as the master plan for IS activities, and IS strategy as the shared view of IS within the organization (Chen et al., 2010, p. 239). In line with these different conceptualizations, (Melville et al., 2004) employ a resource-based-view (RBV) perspective in order to understand how IT is related to organizational performance. They find that while IT can offer a range of benefits for organizations, “the high degree of complexity leads to a context-contingent set of synergistic combinations of IT and other organizational resources...” (Melville et al., 2004, p. 311). In other words, the trick is to choose the right system or combination of systems for a specific context, instead of falling for hypes and simplified models.

(Henderson & Venkatraman, 1993) presents the concept of strategic alignment, triggering a whole stream of research that investigates the question of how to align an organizations’ IS strategy with its business strategy. They state that IS and business strategy have to be linked and aligned on both external (strategic) and internal (operational) levels (Henderson & Venkatraman, 1993, p. 476) and propose 4 different alignment perspectives, depending on the context and the question if business or IS strategy is the driving force for the overall strategy of a company (Henderson & Venkatraman, 1993,

pp. 477–479). Researchers have investigated strategic alignment from various theoretic perspectives and added to the knowledge base in terms of influencing variables, complexities and success factors (Kearns & Lederer, 2003; Kearns & Sabherwal, 2007; Reich & Benbasat, 2000; Tallon, 2008).

While SMEs show considerable differences to large firms, a large part of the literature on IS Strategy and strategy alignment is based on and informed by large firms. These firms tend to have substantially higher levels of internal knowledge and skills regarding IS (Margi Levy & Powell, 2005, p. 8). Thus, researchers argue that theories explaining IS Strategy and governance cannot be simply scaled down to be applicable for SMEs since they show “a completely different economic, cultural and managerial environment” (Devos, Van Landeghem, & Deschoolmeester, 2012, p. 206). Based on the fact that IT is generally understudied and small firms are rarely involved in research discussing IT knowledge, some researchers state that the organizational theories (including practices and behavior) that have been developed for large companies “may not be valid in small ones” (Riemenschneider, Harrison, & Mykytyn, 2003, p. 269). The decision process in SMEs is often dominated by the owner or a member of the management board (e.g. chief financial officer) and can be described as reactive and problem-solving oriented (M. Levy & Powell, 2000). The value that strategic IS planning and valuation can bring for SMEs has not caught much attention until very recently. The value of IS investments is often hard to grasp for SME practitioners, due to their explorative and emerging nature. According to (Symons & Walsham, 1988), in order to efficiently evaluate the benefits and costs associated with an IS, one has to fully understand the underlying processes, which may not be the case in high-growth SMEs that have not yet developed a clear organizational structure.

3 METHODOLOGY

3.1 Empirical Setting

Our empirical work builds upon an interpretive, multiple case study methodology grounded on qualitative data (Eisenhardt, 1989). Our choice for an inductive study is justified by the lack of a viable theory that addresses the IS strategy in fast growing SMEs. IS alignment is regarded as a critical factor for the success of businesses (Melville et al., 2004) and yet, the vast empirical literature on the design of IS strategies overlooks to a great extent the challenges that fast growth processes pose (Margi Levy & Powell, 2005). Using a multiple case study approach, we are able to perform cross-case analysis and generate or extend existing theory (Benbasat, Goldstein, & Mead, 1987).

Our sample is composed of ten fast-growing SMEs that operate in Denmark in various sectors: retail, manufacturing, transportation, ICT and financial services. They were screened for introduction in the study based on the annual Gazelle study conducted by the newspaper Børsen in 2012. This study identifies fast-growing companies – named Gazelles – in the country using the following criteria: (1) minimum of 135.000€ in revenues or 67.000€ in gross profit; (2) positive sum of the operating results and (3) doubled turnover/gross profit over the past four years (Børsen, 2012). Firms were further qualified into the sample on the basis of how many times they were awarded the Gazelle prize (e.g. those who had won it several times were favored) and of data accessibility (e.g. in terms of geographical location and willingness to participate in the study).

We argue that Gazelles are particularly interesting as an object of study, since their high growth rates likely put them in extreme situations that lead them to experience needs ahead of market, thereby making them specially sensitive to strategic decisions. More than other companies, Gazelles are heavily depending on leveraging all resources that could help support their growth. Thus, we assume that the influence of decisions concerning both IS strategy and overall business strategy is faster, more visible and more intensive.

3.2 Data Sources

Data was gathered primarily via interviews and supplemented by observations about the companies recorded during the site visits and by secondary sources of information, i.e. the ORBIS database and the website proff.dk. The Chief Executive Officer (CEO) and Chief Information Officer (CIO) were our target informants, as they were deemed to be most knowledgeable about the technological and strategy-related aspects of their companies.

Interviews were semi-structured. We led the questioning, while maintaining flexibility to embrace issues that emerged during the data collection. The interviewees were asked a variety of questions related to their company's organization and strategic use of IT and management systems. All interviews were conducted between March and November 2014. We analyzed the case data using an inductive approach to derive the characteristics of IS strategies in each surveyed organization. After analyzing the cases separately, we conducted a cross-case analysis to locate common themes and compare the differences that emerged across them. Our inferences were grounded on the empirical data provided by our interview transcripts.

4 FINDINGS

4.1 Archetypes of IS Strategies

Table 1 shows an overview of our results with a characterization of the various dimensions of IS strategies in all case companies. As previously mentioned, our point of departure was a broad understanding of IS strategy as encompassing elements of technology, people, processes and networks. Our research is mainly focusing on the IT-valuation of an organization which is defined within its IS strategy. The dimensions in Table 1 were thus not defined a priori, but emerged during the process of data analysis. Our key inference, which is supported by considerable evidence summarized in Table 1, is that companies can be broadly divided into three groups (A,B,C), which represent different archetypes of IS strategy in fast growing SMEs.

Group A encompasses four companies – Art Software, Digital House, InSynch Host and Magazine Electro. While the first three belong to the ICT sector, the latter is an online retail company. These companies have in common a high IT valuation, in that they clearly see IT as their core activity. As the informant at Digital House expressed it:

“IT is very integrated into our way of life”.

They are inherently reliant on IT to a high extent, since their business models are based on and built around IT (in fact three of them sell IT). Interestingly, Magazine Electro, as the only non-ICT company in the group, presents strong identification with IT, which is reflected in the way it manages all its operations. As the COO commented:

“We try to put as many processes as possible into the IT system (...) One of the areas we can mostly see this is how we deal with our physical chain. We have 45 stores around Denmark that we run from here. It is run by seven people in a small dedicated team. Before we bought it in 2011, it was run by 45 people (...) Because we are focusing on what we can do in the IT way (...) We use the intranet and other IT systems that help and guide the stores around the country in how to run the business. Because we are born e-commerce, IT-driven, this is our way to run a physical chain”.

Characterization of IS Strategies Among Case Companies

| Group | Company* | How IT is seen | IT Reliance | IT Sourcing | IT Responsible | Strategic Planning | IS | IT Budget | IT Innovation Strategy |
|-------|------------------|--|-------------|---|--------------------|--------------------|----|--------------|------------------------|
| A | Art Software | IT as the core of the company | High | Combination of insourcing (primarily) and outsourcing (servers) | CTO | Yes | | n.a. | First-mover |
| A | Digital House | IT as the core of the company | High | Insourcing | CIO | Yes | | 200.000 € | Fast Follower |
| A | InSynch Host | IT as the core of the company | High | Insourcing | 3 Vice- Presidents | Yes | | 1.200.000 € | Fast Follower |
| A | Magazine Electro | IT as the core of the company | High | Insourcing | COO | Yes | | 497.000 € | First-mover |
| B | Apolo Sportswear | IT as a support activity | Medium | Outsourcing (combined with the effort to internalize some knowledge) | CEO | Yes | | 700.000 € | First-Mover |
| B | SweetHome | Currently redefining itself as an IT company | High | In the process of insourcing, after several years of complete outsourcing of IT | CIO (being hired) | Yes | | n.a. | First-mover |
| B | North Utility | IT is just everywhere in this company | High | Insourcing | CIO | Yes | | 15.000.000 € | Fast Follower |
| C | LightHouse | IT as a cost | Medium | Outsourcing | CEO + 1 Accountant | No | | 60.000 € | Slow Follower |
| C | Njord Shipping | IT as a cost | Low | Outsourcing | "IT-guy" | No | | 6.700 € | Slow Follower |
| C | Logica | IT as a cost | Medium | Outsourcing (combined with the effort to internalize some knowledge) | IT-manager | No | | 270.000 € | Slow Follower |

Table 1. Case companies and their characteristics.

Characterization of IS Strategies Among Case Companies

| Group | Company* | E-business Skills Demands | Characteristics of the Workforce | Challenges to Growth |
|--------------|------------------|---|--|---|
| A | Art Software | Combination of commercial and technology skills + knowledge of upcoming technological trends (e.g. cloud computing, big data) | High level of involvement and enthusiasm, invest their free time on learning | Did not verbalize any challenges |
| A | Digital House | Interpersonal and communication skills | High level of involvement and enthusiasm, invest their free time on learning | Lack of a clear governance model |
| A | InSynch Host | Combination of commercial and technology skills + knowledge of upcoming technological trends (e.g. cloud computing) | High level of involvement and enthusiasm, invest their free time on learning | Did not verbalize any challenges |
| A | Magazine Electro | Combination of commercial and technology skills + knowledge of upcoming technological trends (e.g. business intelligence) | High level of involvement and enthusiasm, invest their free time on learning | Did not verbalize any challenges |
| B | Apolo Sportswear | Combination of commercial skills and technology education | Divided between digital market experts and operative IT experts | Lack of skilled workers |
| B | SweetHome | Knowledge of upcoming technological trends (e.g. social media, mobile) | Limited IT/IS skills in the process of developing in-house knowledge | Limitations of prior external IT provider |
| B | North Utility | Knowledge of upcoming technological trends (e.g. enterprise architecture, big data) | Enthusiastic IT professionals (some also externally recognized) | IT coordination and vulnerability of existing IT systems |
| C | LightHouse | Limited demand for IT knowledge | Limited IT/IS skills | Lack of IT knowledge to establish a dialogue with suppliers |
| C | Njord Shipping | Limited demand for IT knowledge | Limited IT/IS skills | Lack of IT knowledge to establish a dialogue with suppliers |
| C | Logica | Knowledge of upcoming technological trends (e.g. social media) | Limited IT/IS skills | Lack of manpower and management support for IT projects |

Table 1 (cont.). Case companies and their characteristics.

With regards to governance, group A companies primarily rely on insourcing. That is, they develop in-house most of the IT solutions for own use. For instance, Digital House reported:

“Most of our systems are handcrafted by ourselves. We have fit our IT platforms to our own processes (...) This decision was made because we looked at the way we worked and it didn't actually fit any other software”.

This group of companies has a large internal knowledge base that allows them to build and adapt their own structure and IT governance processes, often informed by standardized models found in IS development, e.g. agile development (Highsmith & Cockburn, 2001) or SCRUM (Deemer, Benefield, Larman, & Vodde, 2012). In terms of internal organization, the companies in group A can be characterized by having a clear delegation of decision power in the figure of the responsible manager(s) who is fully dedicated to IS matters. They are also appointed to handle strategic planning of their systems and have discretion to allocate resources to the IT budget. In addition, these companies are at the forefront of technological trends, being characterized as either first-movers or fast-followers in their innovation strategies¹. As the COO of Magazine Electro indicated:

“Today I don't think that we are very different from our competitors. But every time we are just a little ahead of them in every small area”.

In the words of the InSynch Host's respondent:

“We try to be out at the leading conferences and the technicians come home and say: hey, we have heard that the trend in the US is this. So that usually means that in a year or two it will show up here (...) And then we have a pre-sales department that has a structured way of gathering market requests (...) We simply follow up on what are the trends”.

Another noticeable characteristic of group A is that the employees of these companies are generally highly involved and enthusiastic about their work, as well as willing to dedicate their free time on learning new technologies. According to the respondent at Art Software:

“A lot of our employees can't stop growing their knowledge because they are loving it. That's what they do and that's what they are living for. And they are doing a lot of it in their spare time as well”.

At the same time, group A companies present sophisticated demands for e-business skills², which include employees that can combine commercial and technical skills and that are familiar with upcoming technological trends. As Art Software's manager's said,

“The whole time in this history it is has been difficult to find the right guys (...) Instead of only focusing on having people that are skilled in coding or skilled in business, you have to find people that can understand both”.

¹ First-movers are companies that are typically the first to market a new product, process or service, as they rely on technological breakthroughs from internal R&D efforts. Fast followers are those that focus on incremental changes to products and that use a second-mover strategy to keep risk low. In contrast, slow followers present a more reactive strategy, as they copy and acquire licenses from the technological drivers, while relying on cheap labor or captive market (Ali, 1994).

² E-Business skills are defined as “the capabilities needed to exploit opportunities provided by ICT, notably the Internet, to ensure more efficient and effective performance of different types of organizations, to explore possibilities for new ways of conducting business and organizational processes, and to establish new businesses. e-Business skills are strategic and related in particular to innovation management, rather than technology-management, skills - which are part of ICT practitioner skills” (E-skills for Europe: Towards 2010 and Beyond, 2004).

Finally, three of the group A companies did not verbalize any specific challenges to their growth prospects. The companies that did so (i.e. Digital House) mentioned organization-related issues (e.g. lack of clear governance model) that did not directly point to IS issues.

Group B in turn, involves a more heterogeneous group of companies from various non-ICT sectors – Apolo Sportswear, SweetHome and North Utility. They share similar IS-related problems that have placed them in a transition phase – from seeing IT as a support activity to redefining IT as one of the core activities of the company. They have all been in a process of moving into a higher IT valuation, either because of perceived opportunities or because of the need to overcome IT-related challenges. Some of our findings point to a lack in company identity as the reason for a lack in IS strategy and alignment (Voss, Cable, & Voss, 2006). The manager at SweetHome expressed clearly this transition:

“We are kind of an unusual tech-business (...) [We consider ourselves] more an IT company, but we're actually discussing that at the moment (...) Sometimes if you consider yourself as being only about properties, maybe you [can't see that] your system for managing properties could be used for something else”.

In the case of SweetHome, the transition was prompted by the limitations of the external supplier, to whom the company previously fully outsourced its IT solutions. Yet, due to the high level of growth of SweetHome, the supplier could neither catch up with the company's increasing demand nor maintain a satisfactory level of service. In the case of Apolo Sportswear, the transition was prompted by the opportunity of using more actively social media in marketing, which resulted in a large new digital department. In the case of North Utility, the vulnerability of its IT solutions posed challenges to the design of an efficient enterprise architecture, which led the company to hire expert professionals. As North Utility's CIO reported:

“I need to accept that some things are just beyond my control. So I need to accept the fact that we are doing something that creates code-spaghetti from time to time (...) We also have these big old legacy solutions in excel, and at some point this just does not scalable anymore (...) Sometimes you can say that the foundation is shaking a little bit, we are almost overloading what we have right now”.

In terms of reliance on IT, they report medium to high levels. In the words of the interviewee from North Utility:

“IT is just everywhere in this company”.

Two of the group B companies were in the process of internalizing IT-related knowledge via insourcing practices as to address these challenges. The hiring of IT professionals and dedicated managers was key in this process. As a result, group B companies were moving from a reactive to a more proactive innovation strategy (either as first-movers or fast followers). The transition phase is also reflected in the workforce of these companies, which ranged from employees with very limited IT/IS skills (SweetHome) to enthusiastic and externally recognized ones (North Utility). Besides, the group B companies verbalized their e-business skills demands, which were targeted primarily at knowledge of upcoming technological trends, and also business-related knowledge. As the interviewee at Apolo Sportswear commented:

“It is easy to find people who know about technology, but it is very hard to find people that know about technology and are commercial as well”.

Finally, *Group C* includes companies that operate in various sectors – LightHouse, Njord Shipping and Logica –and that share the same perspective of IT as primarily being a cost factor. They have thus a low valuation of IT. As Logica's IT manager put it very directly:

“IT in our company is traditionally looked as a cost”.

They also outsource most of their IS-related activities to external vendors with a limited effort to internalize IS related knowledge (such an effort was reported only in the case of Logica). The managing director of LightHouse expressed his valuation of IS as the following:

“I think the basic idea of how we build up this organization was to be agile, very slim and to cut off and outsource everything that was not core business. IT and its support is not a core business for us. So I think we will stay with that, I do not know how big the company would need to be to have to hire a guy”.

Nevertheless, they present different degrees of perceived IT reliance. While it is considered high in LightHouse (*“We 100% depend on the digital side or the IT side from morning to dawn”*), it is considered medium in Logica (where integrated ERP system is used for planning, invoicing and resource management), and rather low at Njord Shipping (where standard office solutions e.g. excel are used). None of the three companies reported to have strategic IS planning. Logica’s IT manager commented in this matter:

“We might be even more successful if we were to spend a little more money on clever solutions, but it’s hard work in the company here because there is no clear strategy (...) It’s not a clear strategy that we would switch from our internal hardware here and going all cloud”.

In this regard, only Logica had a dedicated manager to handle IT matters, while the other two companies did not have a structured delegation of such responsibilities, which were managed ad hoc by the local “IT guy” or the CEO. Group C companies are also characterized by being slow followers with regards to their innovation strategy. This is exemplified by the Logica’s marketing activities:

“Just three years ago we were not doing any commercials pointed out for end users, but it has changed now (...) It started with just e-mailing, now we are on linkedin. Right now the discussion is: are we going to facebook or are we not going to facebook?”

Furthermore, group C companies possessed limited IT/IS skills in their workforce. As Njord Shipping’s CEO testified:

“As a smaller company, although we have been growing, we were without skills in the IT area (...) We were offered solutions from various players in that field and we were not able to evaluate and to choose between them, we thought it was hard to understand”.

Two of the group C companies (Njord Shipping and LightHouse) do not verbalize specific demands in terms of e-business skills, signaling that the lack of IT/IS skills is not seen as an issue. This potential misconception can lead to an increased risk and dependency on IT vendors in terms of both service provision and IT-knowledge. Both firms mentioned having difficulties in establishing a dialogue with suppliers as a result of not understanding IT-specific concepts. As LightHouse’s informant stated:

“When we speak to our suppliers, they are so technical (...) It is difficult to always, sometimes to understand what they really mean and if we actually should do this.”

Logica, in turn, expressed clear demands for more knowledge in upcoming technological trends (e.g. social media marketing) and mentioned the lack of manpower as management support for IT projects as important challenges for growth.

4.2 IS Strategy and Growth

Figure 1 depicts a model of growth patterns based on our findings. We will first describe the dimensions and their meaning, before a detailed discussion on the two growth patterns “high IT valuation” and “low IT valuation” is presented. The model follows the categorization in three groups (A, B, C) that is presented and elaborated on earlier in this section.

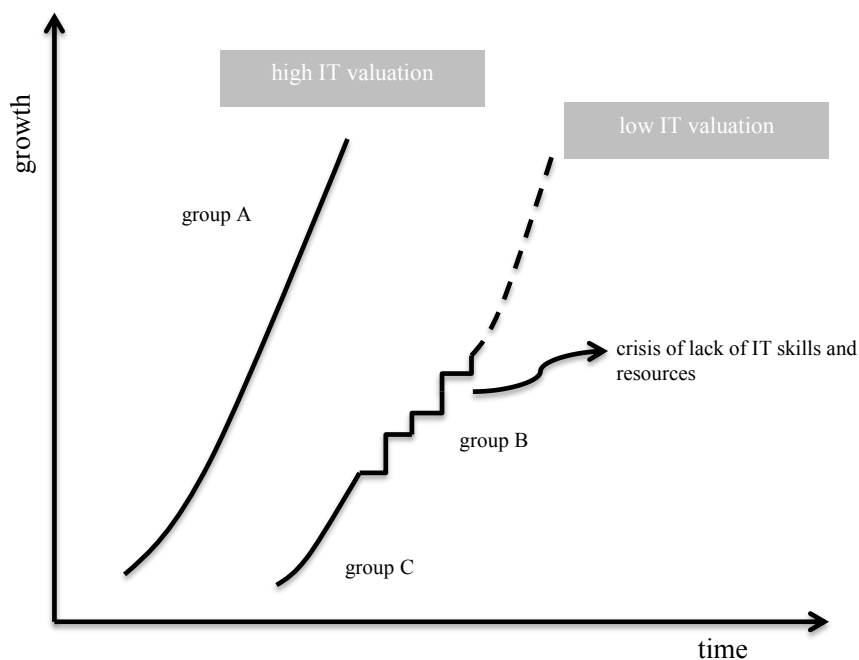


Figure 1. IT valuation of SMEs and growth patterns.

On the horizontal axis the time dimension is represented, while the growth of the organization is shown on the vertical axis. Both dimensions contribute to the change that an organization faces. We argue that it is especially those companies that face high growth in a short period of time (Gazelles) that are most exposed to weaknesses and challenges in their IS strategy, as “problems of coordination and communication magnify, new functions emerge, levels in the hierarchy multiply, and jobs become more interrelated” (Greiner, 1998) when growing.

Figure 1 shows that companies with a high IT valuation (group A) are not facing any IT-related disruptions to their fast growth. None of the group A companies voiced IT-related challenges to growth, as can be seen in Table 1. On the contrary, in line with literature (Oh & Pinsonneault, 2007) we can see that a high level of IT-valuation and commitment can act both as support as well as a driver for growth. This is not only true for IT-companies, as exemplified by our findings from the Magazine Electro case. These companies do not report an IT-related growth crisis since they invested timely in scalable IS strategies and established an in-house technical knowledge-base.

The growth pattern labeled “low IT valuation” depicts the growth path for both groups C and B. Companies in group B have a history of undervaluation of their IS strategy. They decided to change their IS strategy either due to IT-related problems for continuous growth (e.g. SweetHome) or in order to realize market potentials (e.g. ApolloSportswear). In these cases, the influence of IS strategy on their future growth is insecure (dotted line) and heavily depending on how much technical debt and lack of knowledge has been accumulated and how this “crisis of lack of IT skills and resources” is managed.

Organizations in group C are characterized by limited IT/IS skills and little IT-valuation. They have not (yet) reached a situation of IT-related growth crisis. There are multiple potential reasons for these companies to either grow successfully without higher IT-valuation or overcome IT-related issues without changing their IS-strategy. First, companies found in group C are in non IT-intensive business areas. Thus, while their lack of IT-valuation is not a cause for growth problems, they are likely to miss out on the opportunities and scalability that come with a sound IS-strategy (e.g. Magazine Electro).

Overall, our model shows that different IS strategies defining the IT-valuation of a company can have substantial positive as well as negative influence on fast growing SMEs. We derived two distinct patterns of growth based on our empirical data that allow us to demonstrate the potential benefits of a sound IS Strategy on organizational growth.

5 CONCLUDING REMARKS

Our study takes a qualitative perspective to analyze the effects of IS strategies on fast growing SMEs. Our key result is the identification of three archetypes of IS strategies which fit different contexts, depending on the challenges of flexibility and scalability faced. The archetypes are tightly connected to the degree of IT reliance of the industry and IT valuation of the company.

The IT companies from Group A were aware of the value that IT delivers for them and present a well thought through IS strategy. These firms are able to fully capitalize on the benefits that IS are offering them. Companies in Group B were in a transition phase and reported benefits in the process of putting more value on IT. In specific, Apolo Sportswear created a large digital department; SweetHome was in the process of insourcing all its development and North Utility was in the process of hiring skilled people. Their motivation was based on either seeing market opportunities (e.g. Apolo Sportswear) or problems experienced with their former IS approach (e.g. SweetHome and North Utility). In contrast, companies of Group C seem to undervalue IT and outsource as much as possible of what they do not consider part of their core business. In doing so they are increasing the risk of creating a knowledge gap that will create considerable higher costs when changing their IS strategy at a later point.

All in all, our findings imply that IS strategy has to be constantly re-evaluated and aligned with business goals (in special growth). The goal should be not to build up a situation where a change or up-scaling becomes expensive because of lack of knowledge or infrastructure. On the contrary, the lack of a depth in IT skills and knowledge can lead to limitations to growth. In contrast to large companies, IS strategies for fast growing SMEs should be less concerned with traditional areas such as Enterprise Application Architecture (EAA) or integration of IS systems, but with creating a flexible IS strategy that supports direct, traceable and immediate capitalization on IS investments.

REFERENCES

- Ali, A. (1994). Pioneering versus incremental innovation: Review and research propositions. *Journal of Product Innovation Management*, 11, 46–61. doi:10.1016/0737-6782(94)90118-X
- Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly*, 11, 369–386. doi:10.2307/248684
- Bernaert, M., Poels, G., Snoeck, M., & De Backer, M. (2014). Enterprise Architecture for Small and Medium-Sized Enterprises: A Starting Point for Bringing EA to SMEs, Based on Adoption Models. In *Information systems for Small and Medium-Sized Enterprises* (pp. 67–96). doi:10.1007/978-3-642-38244-4_4
- Børsen. (2012). *Gazelle Report*. Retrieved from http://img.borsen.dk/img/cms/tuksi4/uploads/img_server_admin.files/1212_filename_2532.pdf
- Chan, Y. E., Huff, S. L., Barclay, D. W., & Copeland, D. G. (1997). Business Strategic Orientation, Information Systems Strategic Orientation, and Strategic Alignment. *Information Systems Research*. doi:10.1287/isre.8.2.125
- Chen, B. D. Q., Worth, F., Preston, D. S., & Teubner, A. (2010). Information Systems Strategy: Reconceptualization, Measurement, and Implications. *MIS Quarterly*, 34(2), 233–259.
- Davis, G. B. (2000). Information systems conceptual foundations: looking backward and forward. In *Organizational and social perspectives on information technology* (pp. 61–82). Springer.
- Deemer, P., Benefield, G., Larman, C., & Vodde, B. (2012). The Scrum Primer v 2.0, 1–20.
- Devos, J., Van Landeghem, H., & Deschoolmeester, D. (2012). Rethinking IT governance for SMEs. *Industrial Management & Data Systems*, 112(2), 206–223. doi:10.1108/02635571211204263
- Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *Academy of Management Review*.
- E-skills for Europe: Towards 2010 and Beyond*. (2004). Retrieved from http://www7.cedefop.europa.eu/etv/Upload/Projects_Networks/Skillsnet/Publications/EskillForum.pdf
- EU. (2005). *The new SME definition: User guide and model declaration*.
- Greiner, L. E. (1998). Evolution and Revolution as Organizations Grow: A company's past has clues for management that are critical to future success. *Harvard Business Review*, 76(3), 55–60. doi:10.1111/j.1741-6248.1997.00397.x
- Henderson, J. C., & Venkatraman, H. (1993). Strategic alignment: Leveraging information technology for transforming organizations. *IBM Systems Journal*. doi:10.1147/sj.382.0472
- Highsmith, J., & Cockburn, A. (2001). *Agile Software Development: The Business of Innovation. Software Management*.
- Kearns, G. S., & Lederer, A. L. (2003). A resource-based view of strategic IT alignment: How knowledge sharing creates competitive advantage. *Decision Sciences*, 34, 1–29. doi:10.1111/1540-5915.02289
- Kearns, G. S., & Sabherwal, R. (2007). Strategic Alignment Between Business and Information Technology: A Knowledge-Based View of Behaviors, Outcome, and Consequences. *Journal of Management Information Systems*. doi:10.2753/MIS0742-1222230306
- Levy, M., & Powell, P. (2000). Information systems strategy for small and medium sized enterprises: an organisational perspective. *The Journal of Strategic Information Systems*. doi:10.1016/S0963-8687(00)00028-7
- Levy, M., & Powell, P. (2005). *Strategies for growth in SMEs: the role of information and information systems*. *International Small Business Journal*. doi:10.1177/026624260602400307
- Levy, M., Powell, P., & Yetton, P. (2002). The dynamics of SME information systems. *Small Business Economics*, 19, 341–354. doi:10.1023/A:1019654030019
- Luftman, J., & Ben-Zvi, T. (2010). Key issues for IT executives 2009: Difficult economy's impact on IT. *MIS Quarterly Executive*, 9(1), 203–213.

- Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information technology and organizational performance: An integrative model of IT business value. *MIS QUARTERLY*, 28, 283–322.
- Muller, P., Gagliardi, D., Caliandro, C., Bohn, N. U., & Klitou, D. (2014). *Annual Report on European SMEs 2013 / 2014 – A Partial and Fragile Recovery*.
- Oh, W., & Pinsonneault, A. (2007). On the Assessment of the Strategic Value of Information Technologies: Conceptual and Analytical Approaches. *MIS Quarterly*, 31, 239–265. doi:Article
- Reich, B. H., & Benbasat, I. (2000). Factors That Influence the Social Dimension of Alignment between Business and Information Technology Objectives. *MIS Quarterly*, 24, 81–113. doi:10.2307/3250980
- Riemenschneider, C. K., Harrison, D. A., & Mykytyn, P. P. (2003). Understanding it adoption decisions in small business: Integrating current theories. *Information and Management*, 40, 269–285. doi:10.1016/S0378-7206(02)00010-1
- Symons, V., & Walsham, G. (1988). The evaluation of information systems: a critique. *Journal of Applied Systems Analysis*, 15, 119–132.
- Tallon, P. P. (2008). A Process-Oriented Perspective on the Alignment of Information Technology and Business Strategy. *Journal of Management Information Systems*. doi:10.2753/MIS0742-122240308
- Voss, Z. G., Cable, D. M., & Voss, G. B. (2006). Organizational Identity and Firm Performance: What Happens When Leaders Disagree About “Who We Are?” *Organization Science*. doi:10.1287/orsc.1060.0218
- Worthen, B. (2007, December 4). The IT Factor: Tech Staff’s Bigger Role. *The Wall Street Journal*. Retrieved from www.wsj.com

PROSPERITY INDICATORS: FOUNDATIONS, CONCERNS AND PROSPECTS OF USAGE IN POLICY MAKING – THE POLICY COMPASS APPROACH

Research in Progress

Ourania, Markaki, National Technical University of Athens, Athens, Greece, omarkaki@epu.ntua.gr

Panagiotis, Kokkinakos, National Technical University of Athens, Athens, Greece, pkok@epu.ntua.gr

Sotiris, Koussouris, National Technical University of Athens, Athens, Greece, skous@epu.ntua.gr

Costas, Koutras, National Technical University of Athens, Athens, Greece, ckoutras@uop.gr

John, Psarras, National Technical University of Athens, Athens, Greece, john@epu.ntua.gr

Abstract

Prosperity indicators aim at making the concept of prosperity measurable, and hold thus the potential to be used for measuring and assessing the impacts of policies and policy measures over the society. Nevertheless, the abstract and multidimensional nature of prosperity, weaknesses around their definition from a methodological perspective and concerns on their legitimacy among other factors pose difficulties on the exploitation of prosperity indicators for public debate. This paper reviews the usage of prosperity indicators in policy making, bringing up developments, trends and concerns around their conceptualization and use. It further presents the Policy Compass approach, which provides an observatory for experimentation with prosperity indicators with the view of influencing the policy-making process at the local and regional level, and thus contributing towards the vision of more responsive governance.

Keywords: Prosperity Indicators, Policy Making, Policy Impact Assessment, Policy Analysis and Evaluation, Open Data, Conference Track: Electronic Government (GOV 2.0)

1 Introduction

With the advent of evidence-based policy making and the growing demand for government accountability of the recent years, it is not hard to observe an obsession with the heavy exploitation of indicators and all sorts of statistical measures in the public policy arena (Jany-Catrice and Marlier, 2013).

The term ‘indicator’ is one that people can easily understand. It is regularly conceived as a sort of ‘statistical measure’ that can adequately capture crucial aspects of a (social) phenomenon to be monitored, particularly when a specific policy measure is enforced to affect it. As put forward by Innes (1989), an indicator is “a set of rules for gathering and organising data, so that they can be assigned meaning”. As far as policy implementation and impact assessment are concerned, an indicator is conceived as a concrete tool for justifying and optimizing resource allocation.

The idea of employing quantitative indicators in order to evaluate policy implementation is not a recent one but goes back to the ‘40s, when the US economy was being evaluated in terms of the ‘Monthly Economic Indicators’ (Wong, 2006). The success of this project was followed by an explosion of indicators for social change in the ‘60s, when actually the term ‘social indicators’ was coined as a means of assessing where we stand and are going with respect to our values and goals, and to evaluate specific programs and determine their impact (Bauer, 1996).

The idea of exploiting social indicators and developing a theory for defining, using, combining and interpreting them, passed gradually from the US Administration to the large international organisations such as the United Nations (UN)¹ and the Organisation for Economic Co-Operation and Development (OECD)². It was no later than the late ‘60s that researchers started to write about the ‘social indicators movement’. As with most scientific and methodological breakthroughs though, the initial explosion of interest and enthusiasm was followed by scepticism and disappointment, which is quite apparent in the articles and studies of the ‘70s. The basic reason for this frustration was the identification of difficulties on the conceptual and the methodological recipes surrounding the definition, calculation and interpretation of indicators. Even more important, the general feeling was that policy makers were almost reluctant to use social indicators, as they did not seem adequate for measuring their concerns on policy evaluation.

Nevertheless, the early ‘90s witnessed a noisy comeback of the ‘indicators methodology’, through the so called ‘community indicators movement’ (Innes and Booher, 2000). This new wave of interest has been significantly motivated by the global questions on environmental matters and has led to a series of approaches, typically associated to the keywords ‘indicators for quality of life’, ‘sustainability indicators’, sometimes combined with other widely used terms in public discourse, such as ‘economic competitiveness’, etc. (Sawicki, 2002).

Although motivated by different concerns, i.e. the demand for social reform in the 60’s and the need for sustainable economic growth and careful resource consumption in the 90’s, these two movements converge under a major common goal, that of ‘improving the living of people and places’ (Wong, 2006). At this point, attention has to be drawn to the fact that one of the major concerns in the indicators construction and exploitation and also one of ‘pitfalls’ identified in the scepticism of the 70’s has been the access to the relevant data and the difficulties in the collection and reliability of the data needed in order to calculate and interpret prosperity metrics. The revolution of the Web 2.0 and the Open Data Movement, conceived as “the idea that certain data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control” (Auer et al., 2007) arguably opens a new arena of experimentation with social indicators.

¹ <http://www.un.org/>

² <http://www.oecd.org/>

Still, the fundamental concerns and the controversy around the methodology of constructing prosperity indicators, as well as on the definition of prosperity itself remain. Additionally, the indicators' causal dimension, namely their use for the interpretation of policy effects, which is the key for turning indicator development into a decision making process seems to be so far neglected (Jany-Catrice and Marlier, 2013).

In this context, this paper attempts to provide an overview of the use of prosperity indicators in policy making. More specifically, departing from the above raised concerns, the paper discusses on the true sense and value of the experimentation with prosperity indicators and presents the Policy Compass approach as a means of bringing the use of the former into a new dimension. The rest of this paper is structured as follows: Section 2 addresses the concept of prosperity and presents some highlights on the "beyond GDP discussion" while also exposing trends and concerns on the design of related indicators. Section 3 presents the Policy Compass approach on the definition and usage of prosperity indicators, while finally Section 4 summarises the issues discussed and the arguments brought forward in this paper and highlights the ways in which the approach in question creates new prospects and advances the use of prosperity indicators in the policy arena.

2 Background

2.1 On the definition of Prosperity

The term 'prosperity' is frequently used as a synonym to other expressions such as 'welfare', 'well-being' or to what is rather abstractly called the 'quality of life', the latter being defined as "*the product of the interplay among social, health, economic and environmental conditions which affect human and social development*" (Ontario Social Development Council).

How to measure prosperity is a very fundamental and interesting question. Well-being or prosperity does not necessarily equal "economic growth". Economic growth means more consumption of goods and services, exchanged for money. Endless growth means endlessly increasing production and endlessly increasing consumption. Social critics have for a long time pointed out the hollowness carried by that thesis, as it is becoming increasingly apparent that infinite growth is impossible on a finite planet.

Many economists or policy-designers may employ a simple indicator such as the Gross Domestic Product (GDP) (Coyle, 2014) to measure prosperity in a country (or society). This is not to be considered strange, as GDP (either in total or in its 'per capita' version) is one of the most successful single indicators in the history of economics, capturing succinctly the value of all goods in the economy, and known to the vast majority of citizens around the world (although, not always with a positive feeling) and to the totality of people working in economics and political science. Yet, problems and shortcomings are also present in GDP and include the unavailability of correct prices for some goods and products (such as state-provided health care), the inability to reflect the technological quality improvements, etc. The basic problem is, succinctly stated, that such a 'dry', single, technical measure fails to reflect a lot of things about everyday life. As put forward by Robert Kennedy in 1968: "*...GDP measures everything ...except that which makes life worthwhile*" (Kennedy, 1968). If real social phenomena, such as the quality of life or societal integration are to be assessed though, one might have to include other indicators, experiment on research assumptions and probably make a smart combination of measures that will reflect the phenomenon targeted.

2.2 Towards new definitions of prosperity - the beyond GDP Discussion

The assumption that the GDP is unable to gauge progress and prosperity is not entirely new. Ever since the concept of sustainable development has emerged, as the notion of development that meets the needs of the present without compromising the ability of future generations to meet their own

needs (WCED, 1987), it has been made clear that environmental and social factors have to be considered as well in measuring prosperity. Throughout the years several attempts have been made internationally to establish appropriate prosperity and sustainable development indicators. In 1993, OECD first published a set of environmental indicators, where the Pressure-State-Response (PSR) framework is used. The latter is based on the concept of causality: human activities exert pressures on the environment that affect its state, i.e. the quality and quantity of natural resources. In response to these pressures, environmental and economic policies are enacted and applied. The latter form a feedback loop to pressures through human activities. In a wider sense, these steps form part of an environmental (policy) cycle which includes problem perception, policy formulation, monitoring and policy evaluation.

The World Bank's attempts for measuring sustainable development led to a new way of thinking on what constitutes wealth and how it could be measured. In the publications "Monitoring Environmental Progress: A Report on Work in Progress" in late 1995 and "Expanding the Measure of Wealth – Indicators of Environmentally Sustainable Development" in 1997, a nation's wealth is determined as the combination of three major capital components, namely produced or human-made assets, natural capital and human resources, the latter including raw labour, human capital, and the elusive, but important, element known as social capital. The dynamics of creating and maintaining wealth are explored through the indicator of genuine saving, i.e. the true rate of a nation's saving after accounting for the depreciation of produced assets, the depletion of natural resources, investments in human capital and the value of global damages from carbon emissions. Negative rates of genuine saving lead eventually to declining well-being.

The need for sustainable development indicators to guide decision making has also been identified by the UN Conference on Environment and Development that took place in 1992. The preliminary draft report that came out in 1996 under the title "Indicators of Sustainable Development: Framework and Methodologies" included 134 indicators, classified under the social, economic, environmental and institutional dimensions of sustainable development, as well as a detailed presentation of the methodology adopted.

The Barometer of Sustainability, developed by Robert Prescott-Allen in 2001 is a two-coordinate measure, designed to gauge human and ecosystem wellbeing together without submerging one in the other. By assigning to both of its axes equal weight, the barometer sends the message that a healthy human sphere and eco-sphere are equally important for achieving sustainable development.

More recently, the Global Footprint Network, an association of researchers and activists founded by the "inventor" of the Ecological Footprint, Mathis Wackernagel in 2003, has been concerned with making the Ecological Footprint indicator popular as a measure for ecological sustainability, promoting its application and refining the method. First conceived in 1990, the Ecological Footprint measures nowadays humanity's demand on the biosphere in terms of the area of biologically productive land and water required to provide the resources we use and to absorb our carbon dioxide emissions.

The 'beyond GDP' discussion is further exemplified by the expert commission established by the German Federal Parliament and the French Government. The latter 'Commission on the Measurement of Economic Performance and Social Progress' has published a series of reports providing validation and legitimacy to the criticism raised about the adequacy of current GDP-based measures of economic performance and their relevance as measures of societal well-being, economic, environmental, and social sustainability.

The activities presented in this section are indicative of the prevalent trend towards a new definition of prosperity and the list of efforts attempting to replace the GDP is long and has resulted in more or less known indicators, analysing various financial, societal and environmental aspects; however none of them has established itself in a similar way as GDP.

2.3 Trends and Concerns in Indicator Design

Following the clutter around the concept and multidimensional nature of prosperity, recent trends on the design and exploitation of indicators in the policy context involve attempts to construct a single, composite, indicator in order to capture a quality-of-life dimension, the use of multiple, separate indicators for social problems, with the aim of capturing single important aspects of everyday life, such as crime rate, poverty level, air pollution, unemployment rate, etc. as well as the preparation of all-inclusive indicator reports, intended for wide distribution and consultation by decision makers and analysts in an iterative fashion.

From a different angle, parallel to international, national and regional, the use of local indicators seems to be gaining ground. The approach, suggested in (Innes and Booher, 2000) draws inspiration from complexity theory and views a city as an evolving organism which grows and adapts to its environment, proposing a layered indicator system which evaluates city performance in terms of system performance indicators that reflect central values of concern to those living in the city and how the whole urban system is working, (ii) policy and programme indicators, providing feedback to policy-makers on how specific programmes and policies are evolving, and (iii) rapid feedback indicators, facilitating citizens, agencies and businesses to make day-to-day decisions.

Putting aside the actual aspects of prosperity captured, initiatives on the development of prosperity indicators have to be further viewed with a certain degree of circumspection. With the discipline of economics historically defining wealth and progress in conjunction with establishing metrics of measurement, it comes as no surprise that the widespread usage of indicators may also have its roots in a fashion of quantification, giving in parallel rise to a trend of putting forth quantified allegations or more specifically numbers as incontestable arguments (Jany-Catrice and Marlier, 2013).

Care should be taken however to avoid turning policy-related discussions in a meaningless collection of statistics or number series, with questionable usability and unidentified direction. There exists a widespread distrust that attempts to measure prosperity-related, abstract concepts, e.g. quality of life, deprivation, welfare, environmental quality, etc. are not always supported by methodologically sound techniques and well-defined policy related frameworks (Innes and Booher, 2000).

The need for a solid formal methodological basis for indicator development appears to be self-evident. Several approaches have been proposed in the literature, each one employing a discrete number of steps, from the early conception of the idea to the final description of the conceived index. The four-step methodology by Coombes and Wong (1994) is indicatively presented here for its simplicity and flexibility. The approach comprises the steps of (i) conceptual consolidation, which corresponds to the process of clarifying the basic concept to be represented by the analysis, (ii) analytical structuring, pertaining to providing an analytical framework within which indicators will be collated and analysed, (iii) indicators' identification, i.e. translation of key factors identified in the previous step into specific measurable indicators and (iv) synthesis of indicators' values, i.e. synthesis of the identified indicators into a composite index (Wong, 2006).

Regardless of the methodological framework adopted, a well-defined and useful indicator should bring together the qualities of policy relevance, analytical soundness and measurability (UNEP, 2014). Additional concerns on the design of prosperity indicators, are raised with regard to the indicators' institutionalization and the avoidance of bias (Wong, 2006), suggesting respectively the need to setup routine procedures to ensure the continuing existence of an indicator and to legitimise the method of the measure, as well as to have them produced by professional statistical agencies with strong awareness of policy issues, but not actual responsibility for them.

These concerns support the argument that indicators should be a matter of experts and thereby come in contrast with the growing trend of re-inventing government on the basis of the Gov 2.0 notion, which emphasizes citizen engagement and dialogue in the design and use of measures of performance and customer satisfaction with government and in the interpretation of these in a complex, changing con-

text (Innes and Booher, 2000), thus advocating the use of indicators to facilitate the work of many players to make better choices, solve problems and be better able to respond to context and change.

In this respect, the fundamental question of how to integrate prosperity indicators in the policy lifecycle remains. At the heart of all policy design and implementation there is a need to understand why policies should be introduced and how well they are working. Prosperity indicators may provide evidence on the outcomes of policies, and thus answers on their expedience and efficacy. Moreover, they can be used to explore patterns of co-variation across different aspects of the phenomena studied (Wong, 2006). However, their analysis stumbles upon genuine difficulties on the identification of the causal relationships between different factors and is at best guided by existing theories or on a priori assumptions. Still, if statistics and indicators are to serve public debate and policy action, the causality dimension has to be addressed as well.

3 The Policy Compass Prosperity Indicator Framework

Along the above lines, controversy around the definition of prosperity, diversity on the scope and scale of application, indicator typology and design methodological issues, legitimacy concerns, and a great multitude of indicator development initiatives and approaches of different interest or even conflicting nature are the characteristics that make up the context in which citizens and the rest of societal actors are called to formulate their own judgement on the status and progress of prosperity in their countries or communities. Apparently, an ICT tool that allows the former stakeholders to experiment with and leverage prosperity indicators as an aid for policy analysis and evaluation becomes particularly important in this context.

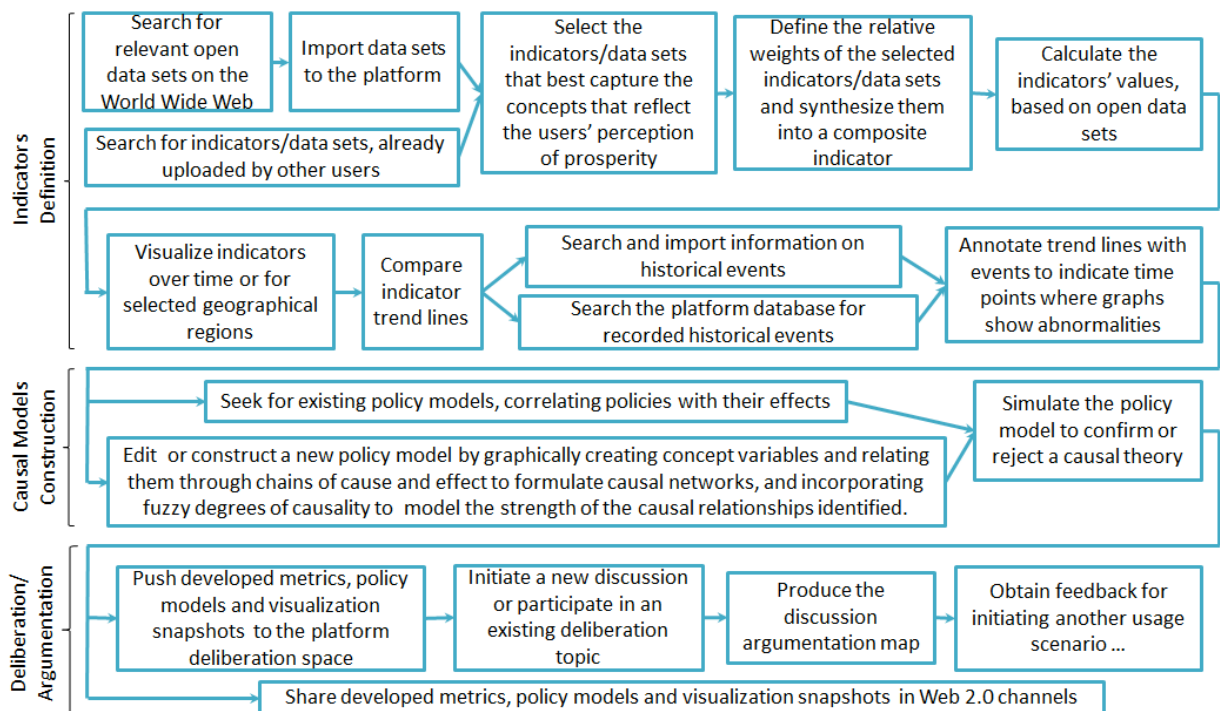


Figure 1. Policy Compass condensed approach

Targeting overall to facilitate factual, evidence-based, transparent and accountable policy evaluation and analysis, and following the view that "Indicators and performance measures have become an im-

portant element in policy initiatives relating to sustainability and to the re-invention of government” (Innes and Booher, 2000), the Policy Compass FP7 project³ places particular emphasis on the design and usage of prosperity indicators. The project aspires more specifically to open new possibilities for the user to correlate, annotate and visualize data from various sources, with an emphasis on the ultimate aim of using them for evaluating policies at the local or regional level through the definition and monitoring of appropriate indicators. To this end, it provides a user-friendly, web-based platform, interfacing public open data sources and integrating a number of toolsets, that implement among others visualization, annotation, causal models’ development, deliberation and argumentation functionalities, in order to offer citizens and societal actors a broad field for experimentation with prosperity indicators, overcoming the challenges and limitations, imposed by current conditions. On top of that, it incorporates web 2.0 features, thus offering the aforementioned stakeholders the capability to share information and engage in meaningful discussions.

Setting out from the citizens’ or other interested stakeholders’ wish to analyse and monitor the effectiveness and efficiency of political actions, a quite comprehensive scenario around the usage of prosperity indicators encompasses, as illustrated in Figure 1, the activities of indicators’ definition, causal policy models creation in the interest of causally analysing changes in the prosperity level, and online deliberation and argument mapping, while simpler use case scenarios may employ selected steps of this workflow.

Currently under development and soon to be released as an operational prototype, the Policy Compass platform, and thereby the respective approach is to be evaluated through real case pilot scenarios at both regional and local level. More detailed information on the Policy Compass project and framework may be found in (Markaki et al., 2014a; 2014b). In brief, the strength of this powerful, Gov 2.0-oriented approach lies in allowing individuals to leverage the data released by governments and public organizations around the world and to create value by contributing their own views and perceptions of prosperity. Ultimately, citizens can formulate informed judgements, hold governments accountable and thus actively engage in policy design.

4 Discussion and Conclusions

The aim of this paper has been to provide a review of the important and fast-growing field of the use of prosperity indicators in policy modelling with the view to synthesize the background and the rationale that has led to the development of the Policy Compass approach and to bring to light the lessons learnt along the course of the project. In the interests of the completeness, the main issues addressed in this paper with regard to the indicators’ applicability and legitimacy as well as the corresponding assumptions are summarised in the following paragraphs.

Experimenting with indicators that may complement or challenge GDP, seems completely feasible, technically speaking, yet it may not be “scientifically legitimate” at all times. In the literature, one can find strong statements in favour of the position that it makes little sense to propose alternatives to established and widely accepted and mature indicators, such as GDP. The latter has achieved a considerable level of legitimacy in quantitative representations and judgements of what wealth really is at the macroscopic level; thereby it seems very ambitious to try to replace it. However, the deep questions on ‘what is a prosperous society’ or a ‘flourishing city/region’, are put forward again and again. The purely economic response based on single indicators such as the GDP is consistently questioned (Jany-Catrice and Marlier, 2013). In particular, there is a growing concern on the limits of GDP-like measures for measuring societal welfare.

³ Policy Compass | Opening Policy Horizons: Open Data-Driven Analysis and Impact Evaluation, <http://policycompass.eu/>

On the other hand, it makes perfect sense to construct aggregate, headline or composite indicators, in order to measure important social phenomena at the regional or municipality level. A strictly scientific approach can be pursued but also experimental approaches, with empirical weighting and evaluation schemes, are welcomed by the community of policy theoreticians (Jany-Catrice and Marlier, 2013). The major question here is of course, the ‘legitimacy’ of the proposed indicator, as it may be considered ad hoc or arbitrary in a subsequent political debate. This is an interesting and largely unexplored issue, which has to do also with political perspective.

Nevertheless, attention has to be drawn to the fact that there do not exist ‘objective’ or ‘neutral’ indicators of any kind. The community does not pursue objective indicators; it rather attempts to construct useful ones. Experience shows that, at the local or regional level community participation should involve the peaceful co-existence of economic, social, and environmental goals around some general vision of well-being, and a vision for the future. The construction and monitoring of the indicators should preferably be a community participation process and concern the setting of goals or benchmarks for monitoring progress of conventional policy along with social capital. Besides, the overall environment is highly encouraging: the advent of the open data movement and Web 2.0 provides a fantastic observatory of experimentation with social indicators for policy evaluation and the strengthening of the democratic process in societies. In this context, it should be clear that a broad experimentation on the calculation and exploitation of indicators, given the multitude of data available, is definitely interesting scientifically and will provide useful feedback.

The Policy Compass prosperity indicator framework is well in accordance with this perspective, as it intends to contribute to, reinforce and promote research and practice towards employing quantitative techniques to circumscribe social phenomena and thus evaluate the results of planned or enforced policy measures along the following key axes:

- Exploring the limits of social computing with quantitative indicators for policy design and assessment, given the unprecedented access to meaningful data provided by the open data sources available.
- Enhancing the experimentation with various kinds of social indicators, ranging from well-known and widely established metrics, such as the GDP and its variants, to the composite and headline indicators which can suitably apply to the regional or municipality level.
- Experimenting on the cross-fertilisation of today’s ICT capabilities with the ideas and intuition of the social (and political/economic) motivation of describing societal welfare with well-defined, representative metrics.
- Importing the cause/effect component of policy analysis, directly into the indicator analysis process through the ‘injection’ of causal analysis tools in the Policy Compass platform and methodology.

Altogether, the Policy Compass methodological framework is anticipated to enrich significantly the experience of using prosperity metrics in the policy arena. Primary evidence on that prospect is going to come from the experimentation taking place in the frame of the project’s pilot applications, so scheduled future work within the project shall include and put emphasis on the analysis of the pilot findings with regard to potential improvements to the project methodology, the project platform and the levels of acceptance of the proposed approach from citizens and other societal actors.

References

- Auer, S.R., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak R. and Z. Ives (2007). "A Nucleus for a Web of Open Data." *The Semantic Web. Lecture Notes in Computer Science* 4825, p. 722.
- Bauer, R.A. (1996). *Social Indicators*. MIT Press.
- Commission on the Measurement of Economic Performance and Social Progress. URL: <http://www.stiglitz-sen-fitoussi.fr>
- Coombes, M. and C. Wong (1994). "Methodological steps in the development of multi-variate indexes for urban and re-gional policy analysis." *Environment and Planning A* 26(1), 297-316.
- Coyle, D. (2014). *GDP: A Brief but Affectionate History*. Princeton University Press.
- Global Footprint Network. *Advancing the Science of Sustainability*. URL: <http://www.footprintnetwork.org/en/index.php/GFN/>
- Innes, J.E. (1989). *Knowledge and Public Policy: The Search of Meaningful Indicators*, 2nd Expanded Edition. New Brunswick, NJ: Transaction Publishers.
- Innes, J.E. and Booher (2000). "Indicators for Sustainable Communities: A Strategy Building on Complexity Theory and Distributed Intelligence." *Planning Theory and Practice* 1(2), 173-186.
- Jany-Catrice, F. and G. Marlier (2013). "Regional Indicators of Well-Being: The Case of France." In: *Community Quality-of-Life Indicators: Best Cases VI*. Ed. By M.J. Sirgy, R. Phillips, D. Rahtz, Springer, p. 19-44.
- Kennedy, R.F. (1968). "Remarks at the University of Kansas." Speech, March 18, 1968. URL: <http://www.jfklibrary.org/Research/Research-Aids/Ready-Reference/RFK-Speeches/Remarks-of-Robert-F-Kennedy-at-the-University-of-Kansas-March-18-1968.aspx>.
- Markaki, O., Kokkinakos, P., Koussouris, S., Löhe, M. and Y. Glikman (2014a). Evidence-based, Transparent and Accountable Policy Analysis and Evaluation: The Policy Compass Approach. In: *Proceedings of Transforming Government Workshop, 2014*.
- Markaki, O., Kokkinakos, P., Koussouris, S., Glikman, Y. and H. Lee (2014b). Fusing Open Public Data, Prosperity Indexes, Fuzzy Cognitive Maps and Argumentation Technology for more factual, evidence-based and accountable policy analysis and evaluation. *Innovation and the Public Sector* 21 (2014), 175 – 184.
- Ontario Social Development Council, see definitions and measures at <http://www.gdrc.org/uem/qol-define.html>.
- Organisation for Economic Co-operation and Development, OECD (1993). *Environment Monographs N° 83 - Core Set of Indicators for Environmental Performance Reviews*. Paris.
- Prescott-Allen, R. (2001). *The Wellbeing of Nations – A Country by Country Index of Quality of Life and the Environment*. Island Press.
- Sawicki, D.S. (2002). "Improving community indicator systems: injecting more social science into the folk movement." *Planning Theory & Practice* 3(1), 13-32.
- United Nations (1996). *Indicators of Sustainable Development: Framework and Methodologies*. New York.
- United Nations Environment Programme, UNEP (2014). *Green Economy - Using Indicators for Green Economy Policymaking*.
- Wong, C. (2006). *Indicators for Urban and Regional Planning: the interplay of policy and methods*, 1st Edition. London and New York: Routledge.
- World Bank (1995). *Monitoring Environmental Progress: A Report on Work in Progress*. September 1995.
- World Bank (1997). *Expanding the Measure of Wealth – Indicators of Environmentally Sustainable Development*. June 1997.
- World Commission on Environment and Development, WCED (1987). *Report of the World Commission on Environment and Development: Our Common Future*. Oxford: Oxford University Press.

USING SOCIAL MEDIA MONITORING FOR PUBLIC POLICY MAKING – AN EVALUATION

Complete Research

Androutsopoulou, Aggeliki, University of Aegean, Samos, Greece, ag.andr@aegean.gr

Charalabidis, Yannis, University of Aegean, Samos, Greece, Greece, yannisx@aegean.gr

Loukis, Euripidis, University of Aegean, Samos, Greece, eloukis@aegean.gr

Abstract

Social media monitoring has been initially adopted by private sector firms in order to collect opinions, complaints and questions concerning their products and services, to be used for making appropriate changes and improvements of them and also for designing communication strategies. Recently government agencies have started adopting SMM, as a form of 'passive citizen-sourcing', in order to collect useful information from citizens concerning their needs, problems, opinions and suggestions, to be used for public policy formulation. It is therefore important to evaluate these first initiatives, so that the potential of SMM with respect to public policy making can be exploited, and at the same time appropriate adaptations and improvements of relevant ICT platforms and practices can be made, in order to reach higher levels of maturity. This paper makes a two-fold contribution in this direction. Initially it develops a framework for evaluating the use of SMM for supporting policy making, initially from the 'classical' ease of use perspective, and then from a public policy perspective, based on the wicked social problems theory. This framework is then used for the evaluation of three pilot applications of a novel method of SMM by government agencies and other policy stakeholders, which has been developed as part of a European research project.

Keywords: Social Media Monitoring, Government, Public Policy, Wicked Problems, Crowdsourcing.

1 Introduction

Social media monitoring (SMM) has been initially adopted by private sector firms in order to collect opinions, complaints and questions that have been posted in various social media (e.g. forums, blogs, Twitter, Facebook, news feeds, etc.) about their products and services, and also about the ones of their competitors, which are used for making changes and improvements of their products and services, for designing communication strategies, and for making appropriate communication interventions in the social media (e.g. replying to negative postings and questions) (Croll & Power, 2009; Sen, 2011; Kasper & Kett, 2011; Järvinen et al., 2015). Government agencies have been traditionally monitoring citizens' opinions and attitudes towards their policies and activities (e.g. through surveys based on representative citizens' samples), and also relevant articles in the 'traditional' media (e.g. newspapers). The emergence and wide use of social media by citizens provide an efficient channel for the creation and exchange of extensive political content, and also for the quick organization of collective political action by citizens (Chadwick, 2009; Soares and Joia, 2015); the political potential of social media poses serious challenges to government agencies, as 'strategic surprises' can suddenly emerge due to the rapid expansion of issues and also the mobilisation and synchronization of numerous citizens enabled by the social media (Shirky, 2011; Bekkers, Edwards, Moody & Beunders, 2011). For the above reasons government agencies have recently started adopting SMM, as a form of 'passive citizen-sourcing', in order to collect useful information from the citizens about their needs, problems, opinions and suggestions, to be used for public policy formulation (both for the design of new policies and for the improvement of existing ones), and also for the design of relevant communication strategies (Bekkers et al, 2013; Charalabidis et al., 2014). This constitutes a big innovation in the policy formulation practices and processes of government agencies, which can help them to address the increasing complexity and 'wicked' nature of the problems of modern societies (Rittel and Weber, 1973; Conklin and Begeman 1989; Conclin 2003).

It is therefore important to evaluate these first initiatives of using SMM in government from various perspectives, in order to learn from them as much as possible, and develop new knowledge in this recently emerged area. Their evaluation will enable us to understand better the value and benefits that SMM can offer to government agencies, so that its potential can be exploited to the highest possible extent, and at the same time will allow us to identify its weaknesses and critical success factors, so that appropriate adaptations and improvements of relevant ICT platforms and practices can be made, in order to reach higher levels of maturity in this area. This paper makes a two-fold contribution in this direction:

- i) It develops a framework for the evaluation of the use of SMM from a public policy making perspective, assessing to what extent it is useful for addressing the fundamental complexities of public policy formulation, based on sound theoretical foundations (wicked problems theory – see section 2.2);
- ii) It uses this framework for the evaluation of three pilot applications of a novel method of SMM by government agencies and other public policy stakeholders (described in more detail in Charalabidis et al. (2014); Loukis and Charalabidis (2014)), which has been developed as part of the European research project NOMAD ("Policy Formulation and Validation through Non-moderated Crowdsourcing" – for more details see www.nomad-project.eu/), partially funded by the "ICT for governance and policy modeling" research initiative of the European Commission.

The paper is structured in seven sections. In the following section 2 the background of our research is presented. The proposed evaluation framework is presented in section 3. The abovementioned novel method we have developed for SMM in government is outlined in section 4, while the research method of our study is described in section 5. Then results are presented in section 6. Finally, in section 7 the conclusions are summarized and future research directions are proposed.

2 Background

2.1 Social Media Monitoring in Government

Social media were initially used by private sector firms, mainly in their marketing and customer service activities, and later were adopted and utilised by government agencies as well, in order to take advantage of the large numbers of users that social media attract, and the unprecedented capabilities they provide to simple non-professional users for developing, distributing, accessing and rating/commenting various types of digital content, and also for the creation of on-line communities (Chun et al. 2010; Bertot et al. 2012; Bonsón et al. 2012; Chun et al. 2012; Margo 2012; ; Nam, 2012; Criado et al. 2013; Ferro et al., 2013; Klischewski, 2014; Zheng and Zheng, 2014; Stamati et al., 2015; Wahid and Sæbø, 2015).

The first generation of social media exploitation by government agencies focused on setting up and operating their own accounts in several social media, posting policy-related content to them (in order to stimulate relevant discussions with citizens), and then analysing citizens' interactions with this content (e.g. views, likes, retransmissions, textual comments, etc.), either manually, or automatically, using central systems accessing various social media platforms through the application programming interfaces (APIs) of the latter (Charalabidis and Loukis, 2012; Ferro et al., 2013; Loukis et al., 2014). Recently a second generation of social media exploitation by government agencies has emerged, which focuses on the collection and exploitation of policy-related content created by citizens beyond their own social media accounts, in various political forums, blogs, news websites, and also in various 'external' Twitter, Facebook, etc. accounts, through SMM.

SMM, defined as 'the continuous systematic observation and analysis of social media networks and social communities' (Fensel et al., 2012), as mentioned in the Introduction, has been initially used by private sector firms, in order to address their fundamental need for listening to their existing and potential customers (in a better and more efficient way than the traditional methods used for this purpose, such as questionnaire surveys), as well as to harness and exploit the wealth of user-generated content available online (Croll & Power, 2009; Sen, 2011; Kasper & Kett, 2011; Fensel et al, 2012; Stavrakantonakis et al., 2012). This is usually conducted through specialised ICT platforms, which enable listening to social media users, and accessing real customers' opinions, complaints and questions, at real time in a highly scalable way, and then measuring and analysing their activities and content concerning a specific brand, or an enterprise, or specific products and services, and processing this information; this leads to valuable insights from the side of enterprises regarding how customers view them, their services and solutions, and also their competitors, and provides support for the design of relevant strategies.

However, there is a lack of frameworks for the multi-dimensional evaluation of SMM platforms, practices and approaches in general, which would allow assessing various aspects of them, and identifying their strengths and weaknesses; this would be very useful for the wider diffusion of SMM, as it would provide evidence for the value and benefits it can offer, and at the same time support for its its improvement. There is only a framework for evaluating SMM tools proposed by Stavrakantonakis et al. (2012), which comprises a set of evaluation criteria that can be used to analyze and assess the functionality of social monitoring ICT tools from three perspectives: the concepts they implement and support (data capture and analysis, workflow, engagement – reaction to posts, and identification of influencers), the technologies used (listening grid adjustment, near real-time processing, integration with third party applications, sentiment analysis, historical data) and the user interface they provide (dashboard, results' export).

Quite limited is the previous literature concerning the use of SMM by government agencies. Bekkers et al. (2013) investigate the practices of SMM in four Dutch public organizations (the Ministry of Education, Culture, and Science, the Ministry of Infrastructure and Environment, the Dutch Tax and Customs Administration, and the Employee Insurances Agency). They examine the goals of SMM (fine-

tuning existing policies, on-line reputation management and prevention of resistances, policies co-production), the way of its operation and its effects. Also, they discriminate between four types of monitored citizens' electronic discussion media based on two criteria: the level of perceived privacy (low or high), and the type of issues discussed (personal or societal); they recommend that more ethical questions arise, so government agencies should be more careful and also transparent, if the citizens' electronic discussion media monitored are characterised by higher perceived privacy and host discussions on more personal issues. Therefore further research is required towards the evaluation of the use of SMM by government agencies from various perspectives; also, there is a lack of a framework for this purpose. Our research contributes to filling these research gaps.

2.2 Wicked Social Problems Theory

Previous literature has analysed the inherent high complexity of the problems of modern societies that have to be addressed by government agencies through appropriate public policies. In a highly influential paper Rittel and Weber (1973) theorize that social problems have become after the end of the second world war 'wicked', as they lack clear and widely agreed definition and objectives. Because our societies have become more heterogeneous and pluralistic in terms of culture, values, concerns and lifestyles, social problems have many stakeholder groups with different and heterogeneous problem views, perceived issues and problems, concerns and expectations, so there is a lack of clear and widely agreed problems' definitions and objectives, and also quite different attitudes to existing policies or proposed policy options.

This increases significantly the difficulty and complexity of public policy formulation, as the responsible for them government agencies have to collect a variety of issues perceived by different stakeholder groups with respect to a policy under discussion, and also their different concerns, expectations and attitudes, and then try to have consultations and negotiations with them in order to achieve some degree of synthesis and consensus. These can be greatly supported through the use of appropriate information systems, referred to as 'issue-based information systems' (IBIS), which allow the collection of the multiple issues and problems perceived by various stakeholder groups, and also their heterogeneous concerns, expectations and attitudes (Kunz and Rittel 1979; Conklin and Begeman 1989; Conclin 2003).

Therefore it is important to evaluate the use of SMM by government agencies from this public policy perspective, and assess to what extent it can be useful for addressing the abovementioned inherent complexity of the formulation of public policies for the highly complex wicked problems of modern societies.

3 An Evaluation Framework

Based on the background presented in the previous section a framework has been developed for evaluating the use of SMM in government for supporting public policy formulation. Our framework is at a higher level than the one of Stavrakantonakis et al. (2012), which has been briefly described in section 2.1; while the latter focuses on the provided functionality and user interface by SMM tools, and the technology used, we are focusing mainly on the final outcome of them: the support they finally provide for public policy formulation, addressing its inherent difficulties discussed in section 2.2 (Rittel and Weber, 1973; Kunz and Rittel 1979; Conklin and Begeman 1989; Conclin 2003), and also the effort required for using them and extracting this support.

In particular, according to the technology acceptance research the intention to use a new technology, and also its actual use, are determined mainly by its perceived 'ease of use' (= the degree to which potential users believe that using it would require minimal effort) and its perceived 'usefulness' (= the degree to which potential users believe that using it will enhance their job performance) (Davis, 1989; Venkatesh and Davis, 2000; Schepers and Wetzels, 2007; Hsiao and Yang, 2011). However, relevant literature has stressed that each of these two factors, and especially the latter, should be further elabo-

rated and focused on the particular objectives and specificities of each particular type of information systems (IS) we study. So for developing our evaluation framework we elaborate these two evaluation perspectives based on the particular characteristics of SMM in government. The basic structure of it is shown in Figure 1.

In particular, our first evaluation perspective is (as usual in most IS evaluation frameworks) the ease of use of SMM, assessing both the general ease of use of it, and also the ease of use of its two main components (as described in more detail in the following section 4): the modelling one (enabling the development of a model of the main terms of the specific policy domain, and also the specific public policy, we want to collect relevant content about in the monitored social media), and the results' visualisation one (= how clear and easy to understand are the visualisations of the results).

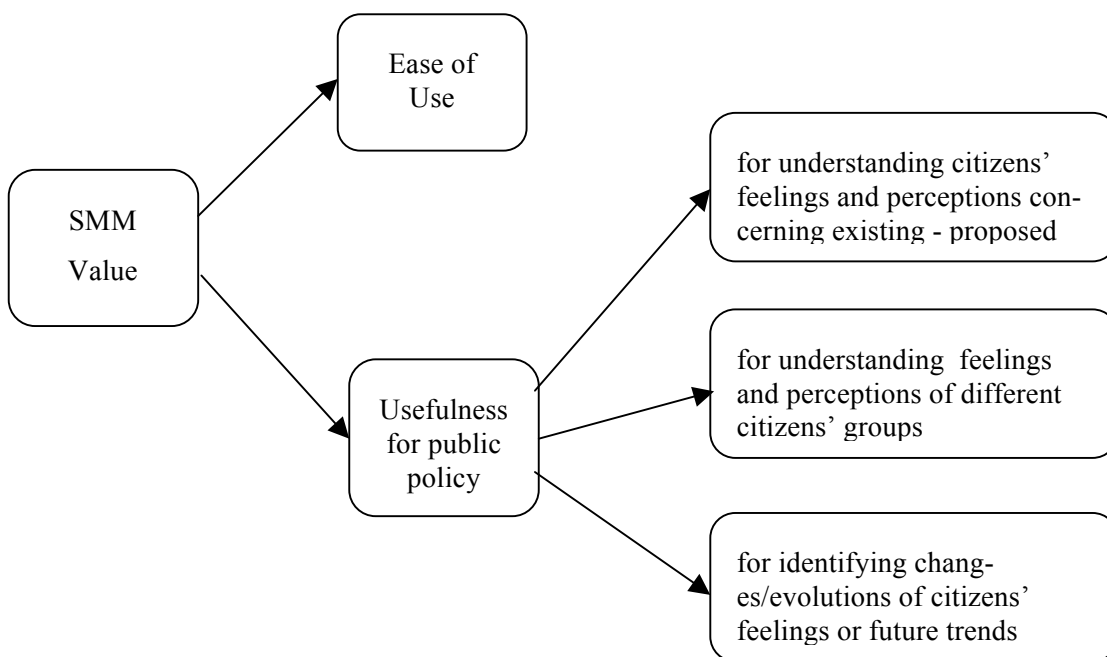


Figure 1. Basic structure of the proposed framework for evaluating SMM use in government

Our second evaluation perspective is the public policy one, assessing to what extent the particular method of SMM in government is useful for the formulation of public policy, and for addressing the inherent complexity of the formulation of public policies for the highly complex wicked problems of the modern highly heterogeneous societies; its theoretical foundation is the wicked social problems theory outlined in the previous section (Rittel and Weber, 1973; Kunz and Rittel 1979; Conklin and Begeman 1989; Conclin 2003). It examines the support provided by SMM for understanding the feelings and perceptions concerning various existing or proposed policies of citizens in general, and also of particular citizens groups (which might differ significantly due to the heterogeneity of modern societies as mentioned above); also, since these are often dynamic, we also examine the support provided by SMM for identifying changes/evolutions in these feelings and perceptions of citizens and relevant future trends. This is associated with the 'dynamic capabilities' (Teece, 2007) of government agencies (concerning their 'sensing' related component). In particular, this evaluation perspective assesses to what extent SMM is useful for the evaluation of citizen's feelings against a prospective or existing policy, or a legislation amendment, and also, going into more detail, of the position of the general public towards different aspects of a suggested policy; furthermore, for evaluating the attitudes of different citizens' groups against a prospective policy, and for the identification of digital opinion leaders (probably associated with important policy stakeholders); and finally, for understanding the timewise

evolution of the public attitude-sentiment against a policy issue/topic, and for the identification of emerging new relevant issues/topics or tendencies in the society. In the following Table 1 the whole evaluation framework is shown.

| |
|--|
| Ease of Use Perspective |
| <ul style="list-style-type: none"> • How easy it is to use this method of SMM in government in general? • How easy it is to use the modelling component ? • How clear and easy to understand are the visualisations of the results ? |
| Public Policy Perspective |
| <p>To what extent this method of SMM in government is useful/beneficial for ...</p> <ul style="list-style-type: none"> • the evaluation of citizens’ feelings against a prospective or existing policy ? • the evaluation of citizens’ feelings against a legislation amendment ? • the evaluation of the position of the general public towards specific aspects of a suggested policy? • the evaluation of the attitudes of different citizens’ groups against a prospective policy ? • the identification of digital opinion leaders ? • understanding the timewise evolution of the public attitude-sentiment against a policy issue/topic? • the identification of emerging new relevant issues/topics in the society ? • the identification of emerging new relevant tendencies in the society ? |

Table 1. A Framework for the Evaluation of SMM Use in Government

4 An Evaluation Framework

A method of SMM use by government agencies, and other public policy stakeholders as well, in order to support the formulation of public policy has been developed as part of the abovementioned European research project NOMAD (Charalabidis et al., 2014; Loukis and Charalabidis, 2014). A brief description of it is provided in this section. It consists of four steps:

i) The first step is to build the ‘domain model’, which is an ontology-based representation of the objects of the “world” (domain) we intend to intervene in through a policy (e.g. energy domain, education domain, health domain). The main entities-terms of its are inserted, as well as relations among them, in a tree structure, using a graphical modelling tool.

ii) Then the second step is to build the ‘policy model’, which is a representation of the public policy we want to collect relevant content about in the social media; it consists of a number of ‘policy statements’ associated with one or more nodes of the policy model, and also for each of them some positive or negative ‘arguments’. A policy model is inserted on a policy model (used as a basis for it) using the above graphical modelling tool.

iii) Upon the completion of the models, the user provides a list of social media sources (e.g. blogs, news websites, and also Twitter, Facebook, etc. accounts) which are going to be crawled, in order to find relevant content about the public policy of interest (= places on the web that according to previous knowledge might contain relevant user-generated content, i.e. citizens are likely to have expressed opinions and suggestions, or relevant needs and problems, concerning this policy topic there).

iv) The defined sources (in step iii) are searched against the above domain and policy models (defined in steps i and ii respectively), and the collected content undergoes sophisticated processing using opinion mining techniques: initially opinions and arguments are extracted, and then sentiment analysis of them is performed (the processing is described in more detail in Charalabidis et al. (2014)). The results are presented to the user in visualised form; a typical results’ visualisation screen (see Figure 2) includes:

- In the upper left part of the screen is shown an estimation of the volume of discussion and the cumulative sentiment for all the elements of the domain or policy model (according to the selections made just above it), the former being visualised through the height of the corresponding rectangle, and the latter through its colour (with the green colour denoting positive sentiment, and the orange negative sentiment).
- For the above model, or for a selected element of it, below (in the lower left part of the screen) is shown the distribution of the volume of discussion over time and also across age groups,
- while in the upper right part is shown a word cloud depicting the most frequent terms-topics discussed online (coloured according to the corresponding sentiment),
- and in the lower left part we can see a list of text excerpts from the sources with relevant content (concerning the selected model or element of it).

Also an 'audience comparative view' can be provided, which shows differences among selected different age, gender or education groups, or differences over time, in the discussed topics (volumes of discussion and sentiment).

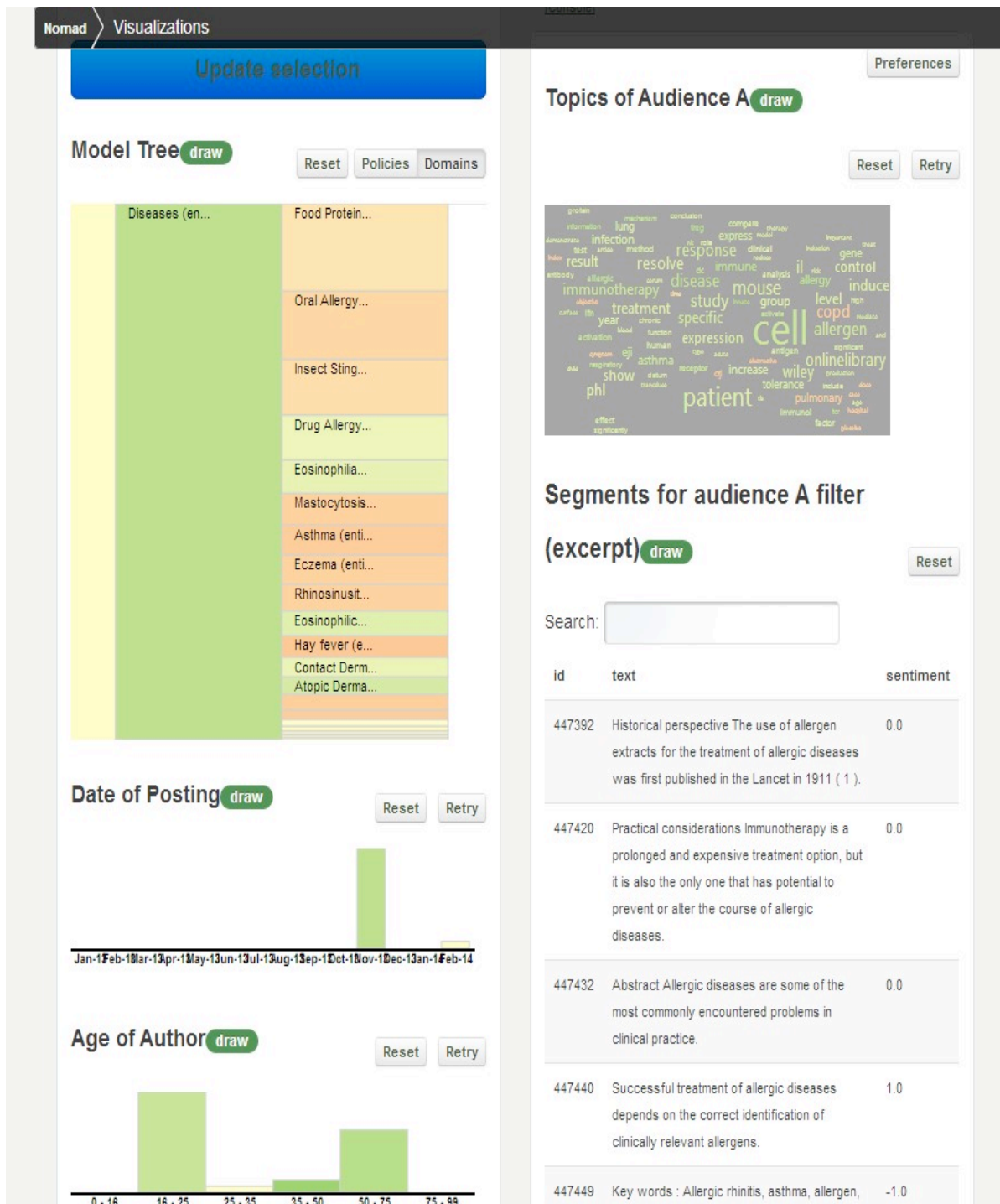


Figure 2. A typical results' visualisation screen

5 Research Method

Three pilot applications of the above method of SMM use in government have been conducted as part of the NOMAD project, and evaluated using the multi-perspective evaluation framework presented above. Since this SMM method, as mentioned in the previous section, is intended to be used not only by government agencies, but also by other public policy stakeholders (who want to know citizens'

opinions, sentiments/attitudes and suggestions concerning various policy related topics before submitting relevant policy proposals to government) as well, two of these pilots were carried out by government organizations: the Greek and the Austrian Parliament; the third one was conducted by an important policy stakeholder in the health domain: the European Academy of Allergy and Clinical Immunology. A detailed scenario has been designed for each pilot, which describes how this SMM method will be used for supporting their policy formulation objectives (focusing on topics reflecting important current debates).

The first pilot application was conducted by the Greek Parliament, and concerns the regulatory and legal framework on energy management, i.e. the “Greek strategy for energy planning” (in compliance with the respective EU Directive 2009/28 EC). The objective of the pilot application was to assess public opinion and attitude/sentiment against this prospective legislation, and based on the collected information through SMM to propose possible amendments. The second pilot application was conducted by the Austrian Parliament, and aimed to monitor the ongoing public debate on the legal basis for “open government information in Austria” and the open government data policies at large. The third pilot application was oriented towards a more scientific policy topic: it was conducted in collaboration with the European Academy of Allergy and Clinical Immunology (EAACI) in order to assist it in discovering the public stance against “allergy diseases and immunotherapy”, and based on this knowledge to design policies for raising awareness in this area, and also to formulate relevant policy proposals.

In particular, for each pilot the following process was followed:

A. At first, the detailed SMM use scenarios in the selected thematic domain have been defined in cooperation with the ‘owners’ of the pilot, and then the domain and policy models required for data crawling were created by them, and finally a list of targeted social media sources (which, according to previous knowledge of the pilot owners, might contain relevant user-generated content) has been specified.

B. After the above preparation, the owners initiated the process of crawling the specified sources against the corresponding domain and policy models, and then processing the collected content, through the ICT infrastructure that supports the application of this SMM method.

C. Then the personnel of the owner organization who participated in this pilot inspected the results, understood them in detail, assisted by members of our research team, and used them in order to draw conclusions about citizens’ opinions, sentiments/attitudes and suggestions concerning the topic of each pilot.

D. Finally, for each pilot an evaluation focus group discussion was organised, which attended by personnel of the owner organization who were involved in this pilot in this pilot, and also other invited persons who had relevant knowledge and experience (e.g. for the pilots of the Greek and Austrian Parliament were invited advisors and assistants of Members of the Parliament, journalists specialise in the corresponding domain; for the EAACI pilot were invited doctors, experts and journalists specialised in allergy and clinical immunology). During this focus group discussion the proposed SMM method was introduced to the audience, together with the supporting ICT infrastructure, and particular applications with their results was showcased. Then the attendees had the opportunity to interact with the ICT platform, performing some predefined tasks, under the observation of organizers’ staff, who supported them in completing these tasks, and recorded any comments or difficulties.

In order to collect evaluation data from the attendees of these focus group discussions we used a combination of both qualitative and quantitative techniques. Qualitative techniques allow a more in-depth examination of about a phenomenon of interest, and therefore the generation of a deeped knowledge about it, not limited to a predefined number of variables, enabling a better and richer understanding of ‘why’ and ‘how’ things happened; on the contrary, quantitative techniques enable condensing and summarizing a large quantity of evidence in a few numbers, which makes it easier to draw conclusions (Maylor and Blackmon, 2005; Ragin and Amoroso, 2011; Yin, 2013). For these reasons in each of these focus groups we conducted initially qualitative discussions focused on the questions of our eval-

uation framework (Table 1), in order to gain a deeper and richer understanding of why the attendees perceive a low or high level of value generated along each of these dimensions. Then we ask them to fill an evaluation questionnaire, which has been structured based on the questions of our evaluation framework: these questions were converted to positive statements, and the respondents were asked to provide the degree of their agreement/disagreement with each of them in a five-levels scale (1 = totally disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = totally agree), which condenses/summarizes all positives and negatives along this value dimension. The above qualitative discussions were recorded with the consent of the participants, and then transcribed and coded manually using an open coding approach (Maylor and Blackmon, 2005); the data collected through the questionnaire were processed using Excel.

6 Results

In Table 2 are shown the results from the processing of the data collected through the questionnaire (average ratings for all questions - evaluation metrics).

| Ease of Use Perspective | |
|---|------|
| It is easy to use this method of SMM in government in general | 3.02 |
| It is easy to use the modelling component | 3.53 |
| It is easy and clear to understand are the visualisations of the results | 3.40 |
| Public Policy Perspective | |
| This method of SMM in government is useful/beneficial for ... | |
| - the evaluation of citizens' feelings against a prospective or existing policy | 4.17 |
| - the evaluation of citizens' feelings against a legislation amendment | 3.69 |
| - the evaluation of the position of the general public towards specific aspects of a suggested policy | 3.94 |
| - the evaluation of the attitudes of different citizens' groups against a prospective policy | 3.40 |
| - the identification of digital opinion leaders | 3.71 |
| - understanding the timewise evolution of the public attitude-sentiment against a policy issue/topic | 4.20 |
| - the identification of emerging new relevant issues/topics in the society | 3.74 |
| - the identification of emerging new relevant tendencies in the society | 3.83 |

Table 2. Average ratings for all evaluation metrics

With respect to the ease of use perspective we can see that though the respondents find that the ease of use of the modelling component and the results' visualizations are moderate to high (average ratings 3.53 and 3.40 respectively), the ease of use of the whole method is moderate (3.02). Therefore the application of this method of SMM in government does not seem to be easy. In the focus group discussions it was mentioned that the main reason for this is the need to build complex models of the specific domain and also the particular policy we are interested in, which requires much time and effort. As a possible solution for this was suggested the use of existing domain ontologies or vocabularies as a basis (and probably add or subtract entities-terms), so the functionality of the ICT platform should be enriched in order to provide such import capabilities. For the results' visualization it was stressed that it is useful for gaining a better understanding of the results, however some improvements are required, such as provision of some additional charts, and improvement of existing ones in order to become more clear and understandable; also it should provide the capability to use some of the results (e.g. terms-topics from the word cloud) in order to improve the domain and policy models. Furthermore, it

was suggested that the visualization tool should be more flexible and adaptable to user's preferences. Another issue raised was that the users cannot understand how the various types of results (e.g. discussion volumes, sentiments, wordclouds) have been produced, and this makes their interpretation difficult; so it would be useful for each chart to provide a basic explanation of how it has been calculated, possibly with links providing more detailed explanations if required by the user (i.e. more results transparency).

With respect to the public policy perspective from Table 2 we can see that the respondents perceive as high to very high the usefulness of this SMM method for evaluating citizens' feelings against a prospective or existing policy in general (4.17), and also between moderate and high (closer to the latter than to the former) its usefulness for making more detailed evaluations of citizens' feelings against a legislation amendment, or of the position of the general public towards more specific aspects of a suggested policy (3.69 and 3.94 respectively). Also, they find moderate to high its usefulness for evaluating of the attitudes of different citizens' groups against a prospective policy, and for identifying digital opinion leaders (which are usually associated with important policy stakeholders) (3.40 and 3.71 respectively). Furthermore, the respondents perceive as high to very high the usefulness of this SMM method for understanding the timewise evolution of the public attitude-sentiment against a policy issue/topic (4.20), and also moderate to high (being closer to the latter than to the former) its usefulness for identifying emerging new relevant issues/topics and emerging tendencies in the society (3.74 and 3.83 respectively).

In the focus group discussions there was an overall agreement that this SMM method provided a time and cost efficient channel to assess citizens' attitudes and feelings on a policy related topic of interest, both from quantity (e.g. the volume of discussion about it) and quality (e.g. the sentiments, the most popular topics within relevant discussions) viewpoint, which is better and less expensive than the traditional citizens' surveys conducted by government agencies. Policy makers or assistants who participated expressed their interest to utilize such SMM results as evidence and documentation for supporting their positions (e.g. on a draft bill, on an existing legislation act, a policy proposal, etc.) "in the parliament or other democratic institutions". Thus, based on their experience in the policy making area the proposed SMM method has the potential to become a "powerful tool for producing new policies", which can be used in all stages of public policies' lifecycle. However, they mentioned the risk of misusing such SMM results for promoting individual interests, by focusing selectively on some of the results that support their own positions, and hiding some others, and possibly misinterpreting them, instead of using these results for really gaining a better understanding of society's opinions, attitudes, sentiments and concerns with respect to an existing policy or a policy under formulation, in order to formulate better, more effective and acceptable policies. Furthermore, they are only to a small to medium extent also mentioned the risk of monitoring citizens' postings perceived by them as private, which would seem to them as an intrusion into their private sphere; even worse would be the use of the results for identifying citizens having political beliefs and orientations different from the ones of government, and for personal monitoring of them. It was generally concluded that the benefits for society from the use of any web-monitoring tool by government depend critically on how this technology is utilised and how its results are exploited.

It has been stressed that one of the most valuable capabilities of this method is the comparative analysis/view it can provide, i.e. present comparisons in the results between demographically different audiences (e.g. in terms of gender, age and education), or different time intervals. The latter comparison between two different time periods enables monitoring the evolution of public stance on a policy related topic, and also measuring the impact and effectiveness of various relevant communication and awareness campaigns or interventions. Participants of the focus groups discussions suggested to include additional location related demographic information, since the geographical dimension is very often important for government decision making, especially for public policies that concern or affect specific regions.

Finally, some of the participants in the focus group discussions mentioned that this SMM method enables to some extent the identification of emerging new relevant issues/topics or tendencies concerning a domain or public policy, however not to the extent they would expect and require. The word cloud does not seem appropriate for the early identification of new issues, topics or tendencies, as it is dominated by the well-established topics-terms (shown with big character sizes, as they are more frequently mentioned by citizens), while the new ones are hardly visible (only some of them are shown with much smaller character sizes, as they are much less frequently mentioned by citizens); new issues, topics or tendencies can be identified mainly by reading the list of text excerpts from the sources with relevant content (lower left part of the typical results' visualisation screen – see Fig.2). In order to have improvement on this two suggestions have been made: a) to add the capability of temporarily removing out of the word cloud the most frequent topics-terms it includes (shown with big size), so that other less frequently mentioned topics-terms become more visible; b) to process further the above text excerpts using various opinion mining techniques, in order to automatically identify new topics-terms.

7 Conclusions

In the previous sections of this paper has been presented initially a framework for the evaluation of the use of SMM by government agencies, from a public policy making perspective (and also from the classical ease of use perspective, which concerns the effort required for its practical application), based on sound theoretical foundations. This framework was then used for the evaluation of three pilot applications of a novel method of SMM by government agencies and other public policy stakeholders, which has been developed as part of the NOMAD European research project.

It has been concluded that such a method of using SMM in government can provide considerable assistance and support for public policy making, as it enables rapid and low cost assessment of citizens' opinions, attitudes and sentiments for a prospective or existing policy, or a legislation amendment; it also allows the identification of differences in the above between different citizens' groups, and also of digital opinion leaders (usually associated with important stakeholders). Furthermore it can provide some assistance and support for understanding the timewise evolution of the public attitude-sentiment against policy issues/topics of interest, and for identifying emerging new relevant issues/topics and tendencies in the society, so it can contribute to improving the 'dynamic capabilities' (Teece, 2007) of government agencies (with respect to their 'sensing' related component). However, this method of SMM does not seem to be easy to use and apply, as it requires building complex models of the specific domain and also the particular policy we are interested in (the use of relevant existing ontologies or vocabularies as a basis for them might reduce the required effort and time for this). Also, the benefits for society from the use of such SMM methods by government depend critically on how and for what purposes they are used, as there are significant risks of misusing them (so transparency in this respect is necessary).

Further research is required concerning the evaluation of SMM in government from multiple perspectives (using different theoretical foundations), and based on a wider range of pilot applications (in various types of government agencies and other policy stakeholders), in order to understand better the value and benefits it can offer to government agencies, and also its weaknesses and critical success factors, so that we can reach higher levels of effectiveness and maturity in this area.

References

- Bekkers, V., Edwards, A. R., Moody, R. and Beunders, H. (2011). "Caught by surprise? Micro-mobilization, new media and the management of strategic surprises". *Public Management Review* 13 (7), 1003–1021.
- Bekkers, V., Edwards, A. and de Kool, D. (2013). "Social media monitoring: Responsive governance in the shadow of surveillance?". *Government Information Quarterly* 30 (4), 335-342.
- Bertot, J. C., Jaeger, P. T. and Hansen, D. (2012). "The impact of policies on government social media usage: Issues, challenges and recommendations". *Government Information Quarterly* 29, 30-40.
- Bonsón, E., Torres, L., Royo, S. and Flores, F. 2012. Local e-government 2.0: Social media and corporate transparency in municipalities, *Government Information Quarterly*, 29, 123–132.
- Charalabidis, Y. and Loukis, E. (2012). "Participative Public Policy Making Through Multiple Social Media Platforms Utilization". *International Journal of Electronic Government Research* 8 (3), 78-97.
- Charalabidis, Y., Loukis, E., Androutsopoulou, A., Karkaletsis and V., Triantafillou, A. (2014). "Passive Crowdsourcing in Government Using Social Media". *Transforming Government: People, Process and Policy* 8 (2), 283-308.
- Chun, S. A., Shulman, S., Sandoval, R. and Hovy, E. (2010). "Government 2.0: Making connections between citizens, data and government". *Information Polity* 15 (1/2), 1-9.
- Chun, S. A. and Luna Reyes, L. F. (2012). "Editorial - Social media in government". *Government Information Quarterly* 29, 441–445.
- Conklin, J. and Begeman, M. (1989). "gIBIS: A tool for all reasons". *Journal of the American Society for Information Science* 40 (3), 200–213.
- Conklin, J. (2003). "Dialog Mapping: Reflections on an Industrial Strength Case Study". In Kirschner, P., Buckingham Shum, P. and Carr C. (Eds.) *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*. London, UK: Springer Verlag.
- Criado, J. I., Sandoval-Almazan, R. and Gil-Garcia, J. R. (2013). "Government innovation through social media". *Government Information Quarterly* 30 (4), 222-230.
- Chadwick, A. 2009. "Web 2.0: New Challenges for the Study of E-Democracy in an Era of Informational Exuberance". *I/S: A Journal of Law and Policy for the Information Society* 5 (1), pp. 9-41.
- Croll, A. and Power, S. (2009). *Complete web monitoring*. Sebastopol: O'Reilly.
- Davis, F. D. (1989). "Perceived usefulness, perceived ease of use, and user acceptance of information technology". *MIS Quarterly* 13(3), 319-339.
- Fensel, D., Leiter, B., and Stavarakantonakis, I. (2012). *Social media monitoring*. Innsbruck: Semantic Technology Institute. URL: <http://oc.sti2.at/sites/default/files/SMM%20Handouts.pdf>
- Ferro, E., Loukis, E., Charalabidis, Y. and Osella, M. (2013). "Policy Making 2.0: From Theory to Practice". *Government Information Quarterly* 30(4), 359–368.
- Järvinen, J., Töllmen, A. and Karjaluoto, H. (2015). "Web Analytics and Social Media Monitoring in Industrial Marketing: Tools for Improving Marketing Communication Measurement". In Leroy Robinson, Jr (Ed) *Marketing Dynamism & Sustainability: Things Change, Things Stay the Same...* New Orleans, USA: Springer Verlag.
- Kasper, H. and Kett, H. (2011). "Social Media Monitoring-Tools". In T. Schwarz (Ed.) *Leitfaden Online-Marketing. Das Wissen der Branche*. Waghausel: Marketing-Borse, 662-669.
- Klischewski, R. (2014). "When virtual reality meets realpolitik: Social media shaping the Arab government–citizen relationship". *Government Information Quarterly* 31(3), 358–364.
- Kunz, W. and Rittel, H. (1979). *Issues as Elements of Information Systems*. Working Paper No. 131, University of California, Berkley.
- Loukis, E. and Charalabidis, Y. (2014). "Active and Passive Crowdsourcing in Government". In: Marijn Janssen, Maria Wimmer and Ameneh Deljoo (Eds) *Policy practice and digital science: Inte-*

- grating complex systems, social simulation and public administration in policy research*'. Springer Verlag – Public Administration and Information Technology Series.
- Loukis, E., Charalabidis, Y. and Androutsopoulou, A. (2014). "An Analysis of Multiple Social Media Consultations in the European Parliament from a Public Policy Perspective". In: *Proceedings of the European Conference on Information Systems (ECIS) 2014*, Tel Aviv, Israel, June 5 - 13.
- Nam, T. (2012). "Suggesting frameworks of citizen-sourcing via Government 2.0". *Government Information Quarterly* 29, 12–20.
- Margo, M. J. (2012). "A Review of Social Media Use in E-Government". *Administrative Sciences* 2 (2), 148-161
- Maylor, H. and Blackmon, K. (2005). *Researching Business and Management*. New York: Palgrave-Macmillan,.
- Ragin, C. and Amoroso, L. (2011). *Constructing Social Research: The Unity and Diversity of Method*. 2nd Edition. California: Pine Forge Press – Sage Publications,.
- Rittel, H. W. J. and Weber, M. M. (1973). "Dilemmas in a General Theory of Planning". *Policy Sciences* 4, 155-169.
- Sen, E. (2011). *Social media monitoring für Unternehmen*. Cologne: Social Media Verlag.
- Shirky, C. (2011). "The political power of social media". *Foreign Affairs* 90, 28–41.
- Schepers, J. and Wetzels, M. (2007). "A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects". *Information & Management* 44, 90-103.
- Soares, C. D. M. and Joia, L. A. (2015). "The Influence of Social Media on Social Movements: An Exploratory Conceptual Model". In: *Proceedings of 7th IFIP 8.5 International Conference ePart 2015*, Thessaloniki, Greece, August 30 – September 2.
- Stamati, T., Papadopoulos and T. Anagnostopoulos, D. (2015). "Social media for openness and accountability in the public sector: Cases in the Greek context". *Government Information Quarterly* 32, 12–29.
- Stavrakantonakis, I., Gagiou, A. E., Kasper, H., Toma, I., and Thalhammer, A. (2012). "An approach for evaluation of social media monitoring tools". In: *Proceedings of the Common Value Management Workshop CVM, co-located with the 9th Extended Semantic Web Conference ESWC2012*, Heraklion, Crete, Greece.
- Teece, D. (2007). "Explicating Dynamic Capabilities: the Nature and Microfoundations of (Sustainable) Enterprise Performance". *Strategic Management Journal* 28, 1319-1350.
- Venkatesh, V. and Davis, F. D. (2000). "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies". *Management Science* 45(2), 186-204.
- Yin, R. (2013). *Case Study Research: Design and Methods*. 5th Edition. California: Sage Publications.
- Wahid, F. and Sæbø, O. (2015). "Affordances and Effects of Promoting eParticipation Through SocialMedia". In: *Proceedings of 7th IFIP 8.5 International Conference ePart 2015*, Thessaloniki, Greece, August 30 – September 2.
- Zhang, B. and Vos, M. (2014). "Social media monitoring: aims, methods, and challenges for international companies". *Corporate Communications: An International Journal* 19 (4), 371-383.
- Zheng, L. and Zheng, T. (2014). "Innovation through social media in the public sector: Information and interactions". *Government Information Quarterly* 31 - Supplement 1, S106–S117.

OPEN AND RESPONSIBLE INNOVATION CONCEPTS FOR COMPETITIVE ADVANTAGE

Research in Progress

Maric, Josip, Montpellier Research in Management (MRM), University of Montpellier, Montpellier, France, josip.maric@etud.univ-montp2.fr

Rodhain, Florence, Montpellier Research in Management (MRM), University of Montpellier, Montpellier, France, florence.rodhain@univ-montp2.fr

Barlette, Yves, Montpellier Research in Management (MRM), Montpellier Business School, Montpellier, France, y.barlette@montpellier-bs.com

Abstract

In this paper we present two evolving concepts, Responsible Innovation and Open Innovation, aiming to deepen discussion on similarities and differences between the two. The idea for this conceptual paper is led by the need to elaborate the two concepts in order to allow easier alignment with organizational strategies. We begin by presenting the definitions of the concepts and contexts in which the notions were developed. Responsible Innovation is seen here as a potentially overarching notion that embeds the characteristics of Open Innovation, so the paper discussion is slightly positioned in favour of Responsible Innovation. Moreover, with this paper we want to make contribution on the two concepts in management and business literature.

Keywords: Management, Responsible, Open, Innovation.

1 Introduction

Innovations were not always responsible, having negative impact on individuals, communities and eco-systems in favour of economic growth and shareholder value creation. Innovation may bring a lot of good to society, but history provides many examples of innovations and new technologies that have had serious negative consequences, or that just failed to make meaningful contributions to society. This puts a pressure to innovate differently in the 21st century. The challenge provides significant opportunities for creating alternative products and services, new business models and ways of working.

It was not until the recent decades that the terms '*responsible innovation*', '*responsible research and innovation*' (Blok et al., 2012) or '*open innovation*' emerged in academic discussions. Terms draw attention of policy makers as well as business organizations in their response to increased attention to social and environmental impact of innovation processes.

Responsible Innovation is a term, a mind-set, developed within EU as a policy tool to tackle grand societal changes. It means involvement of all societal stakeholders (public, private, civic ...) with the innovators. It incorporates responsibility, ethical and moral values in the core process of innovation. As such, it represents governance of research or operational process where innovation's direct and indirect socio-environmental impacts are evaluated throughout the innovation lifecycle.

Open Innovation, on the other side, presents managerial concept in which firms innovate more 'openly', developing new products and services jointly with other partners. Impulse for Open Innovation can come from different group of innovation stakeholders, for instance customers or suppliers, or it can be outcome of cooperation with research institutes or other enterprises.

This paper focuses on fundamental differences and collision points between 'open' and 'responsible' innovation. Although the two terms are contributing each other in several major points, there is a fine, subtle difference that needs to be explored.

Paper is organized as follows: next section analyses available definitions and characteristics of Responsible Innovation. Secondly, Open Innovation is presented. Following section after that is the discussion on the two concepts. Finally, concluding thought of the paper are presented.

2 RESPONSIBLE INNOVATION DEFINITION

In this section of the paper we will examine concept of Responsible Innovation. The guiding assumption of the concept is that "*right from the start, research, development and design [can] incorporate relevant ethical and societal aspects*" (Blok et al., 2012). Ethics in innovation has been discussed by many relevant scholars across disciplines (see in Van den Hooven et al., 2012).

European Commission (EC) workshop in Brussels presented a report in 2011 where Responsible Research and Innovation has emerged on the formal agenda because: "*innovation policy actors across sectors are drawing on past experiences in the hopes of motivating the development of technologies for social benefit, preventing both disasters and the loss of promising technological advances, and addressing public anxiety over unintended and irreversible consequences*", (Guston et al., 2014).

As mentioned previously, Responsible Innovation is a concept developed within EU as a policy tool to tackle grand societal changes. And even though the concept was developed in the domain of the policy making, the researchers who were involved in the development of the concept started taking it one step further. Now there is evolving research community, research centres specialized in the questions of Responsible Innovation and its potential application in the vast number of disciplines. With evolving application scope, questions on the conceptualization and practical acceptability of the Responsible Innovation gain a momentum.

Moreover, there are distinguished streams of researchers who present Responsible Research and Innovation as a theoretical approach of governing the research (or innovation process), while second

stream presents Responsible Innovation as an operational process to anticipate direct and indirect impacts of innovation throughout its lifecycle, as a long term strategy.

Notable promoter of Responsible Research and Innovation is Renè von Schomberg, Dutchman with long term involvement in European Union policy making. According to him, Responsible Research and Innovation marks: “*a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view on the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society)*” (Von Schomberg, 2011).

Xavier Pavie, business scholar in favour of Responsible Innovation as operational process, states: “*Responsible Innovation is an iterative process throughout which the project’s impacts on social, economic and environmental factors are, (...), measured and otherwise taken into account at each step of development of the project, thereby guaranteeing control over, (...), the innovation’s impacts throughout the entire lifecycle*” (Pavie et al., 2013).

In their paper, Pavie and Carthy (2014) discuss about the differences on definitions of evolving research community. As such, study presents that it is of utmost importance to dissociate ‘responsible innovation’ from the concept of ‘responsible research and innovation’ (a central theme in the context of the Horizon 2020 European program presented by Von Schomberg).

Unfortunately, there are still certain barriers concerning Responsible Innovation integration in business strategy. For instance, although well covered, there is still no authoritative definition of Responsible Innovation. But even though, for both described streams of researchers it is common that the idea lies in a more responsible adoption of innovation. The challenge of Responsible Innovation, therefore, lies in a more responsive, adaptive and integrated management of the innovation process.

Opening discussion on the concept is extremely utile in setting up a ground for popularizing Responsible Innovation, as well as existence of different funding programmes and conferences. One of the examples is spark of the Journal of Responsible Innovation, intended to link the concept with various scientific disciplines.

3 OPEN INNOVATION DEFINITION

Open Innovation is a distributed innovation process based on purposively managed knowledge flows across organizational boundaries, using pecuniary and non-pecuniary mechanisms in line with the organization’s business model (see in West et al., 2014). In other words, this means that valuable ideas can come from inside or outside the company and can go to market from inside or outside the company as well.

Open Innovation at the time when it was presented introduced new terminology and offered breakthrough in how enterprises viewed innovation process. Most notable and cited promoter is Henry Chesbrough with his publication from 2003, book on ‘Open Innovation’. For many practitioners, concept of Open Innovation gave a new point of view on Research and Development (R&D), shifting dominant logic of R&D from internal discoveries to external engagements.

Open Innovation has offered a path for considerable impact on practice, primarily among industrial and consulting companies. Examples of use are LEGO, LG or Samsung. Basically, the concept incorporates share of risk among the relevant stakeholders, and as such, this approach spurred scope of new business models for generating revenue.

Future perspective can be seen as summing up debate for Open Innovation assessment, better understanding the role of appropriability and integrating Open Innovation with established theories of management and economics (West et al., 2014).

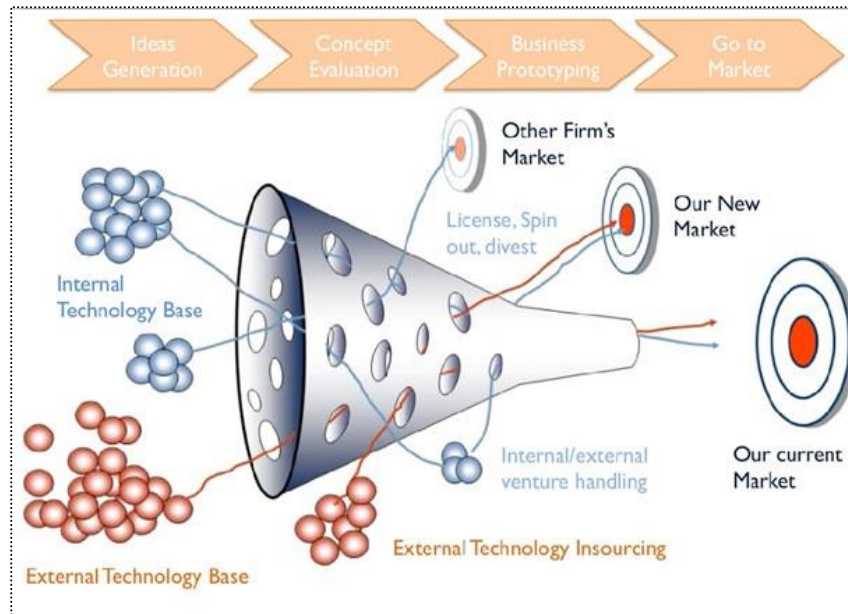


Figure 1. Chesbrough scheme of Open Innovation. (Chesbrough, 2004)

4 DISCUSSION ON THE TWO CONCEPTS

Innovation can be radical or incremental, and it can be applied to products, processes or services and in any organization. It can happen at all levels in an organization, from management teams to departments and on the level of individuals (O'Sullivan, Dooley, 2009).

Since innovation is the central topic both in Responsible Innovation and Open Innovation, can we speak about similarities and differences between the two concepts? Where do Responsible Innovation and Open Innovation intersect, where do they collide and whether they contribute each other? Moreover, do organizations need innovation in their capabilities in order to absorb the full potential of these concepts? There is a considerable overlap between the two concepts and it will be elaborated further in the text.

4.1 Differences

We start listing the differences and similarities on definitions of the concepts. As mentioned, Responsible Innovation presents inclusive approach to innovation, being open to public, private as well as to other (civic) stakeholders. Open innovation promotes openness to relevant stakeholders as well, but unfortunately, there is a gap between the openness towards relevant business partners or to external, university or non-governmental (NGO) partners.

Moreover, Responsible Innovation as a goal has creation of societal value of the innovation, while Open Innovation does not explicitly results in socially desirable outcome of the innovation. Questions on ethically acceptable and environmentally friendly innovation outcome are not priority conditions under Open Innovation concept. This leave space for blurry area of the overall impact of innovation, where hazardous and potentially dangerous outcomes can be neglected in pursue for sole shareholder benefit creation.

4.2 Similarities

Responsible Research and Innovation is a new approach in governing research. Moreover, Open Innovation is linked with theories of governance. As such, core definitions of the two concepts present a collision point. Both concepts incorporate the need for better governance of innovation process, promoting openness to relevant stakeholders.

Long term vision and alignment with organizational business strategy presents second collision point. In the case of Responsible Innovation, it means a proactive organizational response to face uncertainties in business, on a long term basis. Idea is to move towards shared notion of responsibility and to anticipate possible unintended impacts of new technology or process innovations, from the early beginning of innovation process.

On the other hand, Open Innovation is much more firm focused governance of research. Firms chose partial openness strategies to build reputation, gain market share and attract third party contributions (Owen et al., 2013).

Our brief summary of similarities and differences on the two concepts shows there are still open questions that need to be addressed. One is concerning Responsible Innovation, where more work is required on presenting practical benefits for industry partners. Concept still remains in the field of policy makers, and although future holds great promise when discussing this concept, we are witnessing lack of studies on that subject (Pavie, Carthy, 2014; Blaskó et al., 2014).

Moreover, Open Innovation is often misunderstood on how open firms need to be to relevant partners with regards on intellectual property. Open Innovation research still remains linked to interaction between two firms, while a shift to collaboration with external innovation networks, such as universities, NGO's and communication beyond the private sector is potential further development perspective of the concept (see in West, Lakhani, 2008).

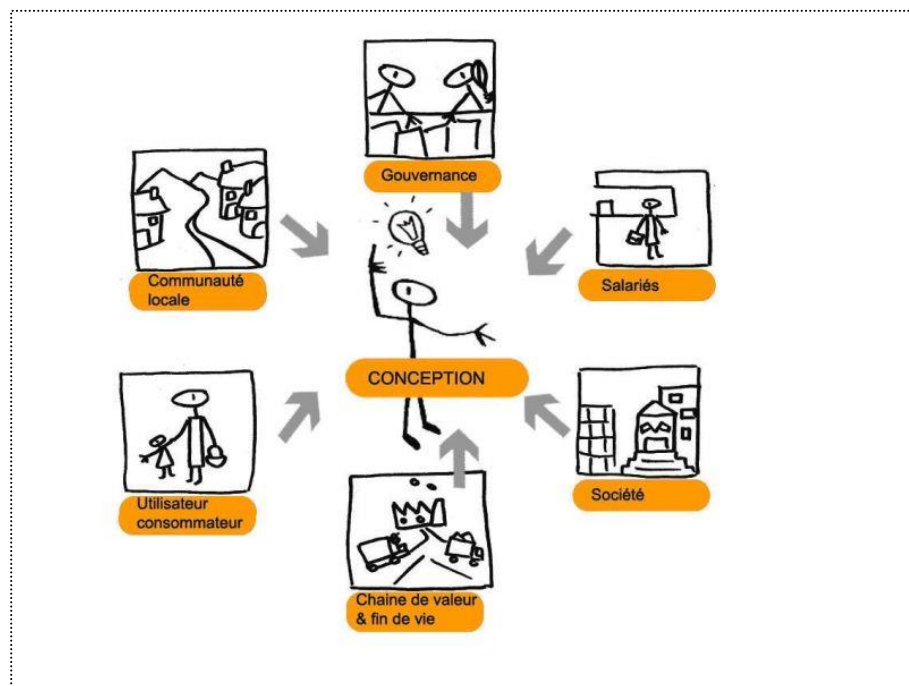


Figure 2. *Involvement of relevant stakeholders in Responsible Innovation is the basis for creating societal value of the innovation. Collaboration of companies with external networks (non-profit organizations) has been sceptically regarded by some Open Innovation scholars. (KARIM Project, 2011)*

4.3 Innovating for innovation concepts?

It is clear that two upper-mentioned concepts, in their cores, carry a notion of major impact on organizational practice when it comes to innovation. But the question is whether organizations have capabilities to innovate differently? Managers increasingly realize importance of developing specific organizational and managerial skills to effectively innovate differently. Unique technical and managerial competencies, along with organizational resources and capabilities require strategic alignment to facilitate concept adoption and knowledge transfer.

Innovating in management principles and processes can create long-lasting advantage and produce dramatic shifts in competitive position (Hamel, 2006). Management innovation has been marked as one of the tools to allow enterprises to achieve a new performance level. Management innovation involves introduction of novelty in an established organization, and as such it represents a particular form of organizational change (Birkinshaw et al., 2008), which would allow absorption of new innovation concepts.

Gary Hamel, scholar focusing on the questions of management innovation, states that companies today are lacking disciplined and organized process for management innovation (Hamel, 2011). Van de Ven (1986) study about problems in the management of innovation clarified four central problems:

- the human problem of managing attention,
- the process problem of managing ideas into good currency,
- the structural problem of managing part-whole relationship,
- the strategic problem of institutional leadership.

Although the innovation field has changed tremendously from the time their study was introduced, characterized by the birth of multiple waves of new innovation management concepts, it is hard to ignore the fact that the Van de Ven problems are still not accurately addressed. This is shown by the fact that 20 years later, Van de Ven and Engleman (2004) still considered them as central in managing corporate entrepreneurship and innovation (see in Vo et al., 2011).

With all of this said, we can state that both, Responsible and Open Innovation, concepts require more work on practical methodology of implementation across organization's managerial levels. Neglecting any further discussion on two relatively new research topics like Responsible or Open Innovation will most likely result in a loss of momentum for both innovation concepts and their adoption with industry sector, losing shared understanding for further development and transfer of knowledge.

5 Conclusion

So far, management in the 21st century isn't much different from management in the 20th century. Therein lies the opportunity for evolution. Forecasting changes in the *business-as-usual*, innovations in management are marked highly desirable to foster product and process innovations (Birkinshaw et al., 2008). In such a case, considering different principles of innovation can come as a handy tool not for short term profit creation, but long term investment in organizational change.

In this paper we have introduced interesting new concepts to govern innovation, and discussed the conditions for innovating responsibly and openly. We have brought definitions of Responsible and Open Innovation and hopefully opened debate on the two concepts.

In brief, both concepts have far-reaching implications on how innovation is performed. It is important to mention that Open Innovation and Responsible Innovation are not mutually exclusive. Hence some future research initiatives could focus on more detailed work of fusion between these two concepts, where companies could be invited to cooperate inside an open-innovation network. Companies require innovation in order to survive and remain profitable in an ever increasingly competitive environment. But becoming a 'responsible' or 'open' business requires integrating a concrete strategy and sustaina-

ble process to that purpose, which cannot happen overnight, and here lies the opportunity for further familiarization with the concepts like Responsible Innovation or Open Innovation.

Acknowledgments

The research work is supported through European Commission EACEA (Education, Audio-visual and Culture Executive Agency) Erasmus Mundus Action Plan 2 External Cooperation Window, under SIGMA program.

References

- Birkinshaw, J., Hamel, G. and Mol, M.J. (2008). Management Innovation. *Academy of Management Review*. 33 (4), 825-845.
- Blaskó, B., Lukovics, M. and Buzás, N. (2014). Good Practices in Responsible Innovation, in a book *Responsible Innovation, in the frame of FaRInn, Facilitating Responsible Innovation in South-Eastern European countries*, pp. 179-192.
- Blok, V., van den Belt, H. and Lemmens, P. (2012). Responsible Innovation? Towards a Critique of the Concept of Innovation, in *Book of abstracts - 2nd Conference on Responsible Innovation, The Hague, The Netherlands*. Retrieved online: www.responsible-innovation.nl
- Chesbrough, H. (2003). *Open Innovation*. Boston: Harvard Business School Press.
- Chesbrough, H. (2004). *Open Innovation; Renewing Growth from Industrial R&D*. 10th Annual Innovation Convergence, Minneapolis.
- Guston, D.H., Fisher, E., Grunwald, A. and Owen, R (2014). Responsible Innovation: motivations for a new Journal. *Journal of Responsible Innovation*. 1 (1), 1-8.
- Hamel, G. (2006). The Why, What and How of Management Innovation. *Harvard Business Review*. 84 (2), 74-84.
- Hamel, G. On the future of management, at the Management Innovation Exchange (MIX), 2011. Available on: <https://www.youtube.com/watch?v=K3-IY66tpI>
- KARIM Project (2011). Responsible Innovation in the context of the KARIM project. Accessed online in July 2014: www.karimnetwork.com
- O'Sullivan, D. and Dooley, L. (2009). *Applying Innovation*. Sage Publications.
- Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E. and Guston, D. H. (2013). A Framework for Responsible Innovation. In R. Owen, M. Heintz, & J. Bessant (Eds.), *Responsible Innovation*. Chichester, UK: Wiley.
- Pavie, X., Scholter, V. and Carthy, D. (2013). *Responsible-Innovation in Practice: How to implement Responsibility across an Organization*. Cahier Innovation & Society.
- Pavie, X. and Carthy, D. (2014). Addressing the Wicked Problem of Responsible Innovation through Design Thinking. ESSEC Working paper. Document de Recherche ESSEC. Accessed online in January 2015. https://hal.archives-ouvertes.fr/file/index/docid/996181/filename/201404_Addressing_the_wicked_problem_of_responsible_innovation_through_design_thinking_Pavie-Carthy_ESSEC_V4.pdf
- Van den Hooven, J., Helbing, D., Pedreschi, D., Domingo-Ferrer, J., Gianotti, F. and Christen, M. (2012). FuturICT—the road towards ethical ICT. *The European Physical Journal Special Topics*. 214 (1), 153-181.
- Van de Ven, A. (1986). Central problems in the management of innovation. *Management Science*. 32, 590-607.
- Van de Ven, A. and Engleman, R. (2004). Central problems in managing corporate innovation and entrepreneurship. *Advances in Entrepreneurship, Firm Emergence and Growth*. 7, 47-72.

- Vo, L.C., Mounoud, E. and Hminda, N. (2011). Central problems in managing open innovation: a recurrent issue? in OLKC 2011: Making waves, Hull University Business School. Retrieved online August 2015. http://www2.warwick.ac.uk/fac/soc/wbs/conf/olkc/archive/olkc6/papers/id_174.pdf.
- Von Schomberg, R. (2011). Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies fields. Brussels: European Commission.
- West, J., Salter, A., Vanhaverbeke, W. and Chesbrough, H. (2014). Open Innovation: The next decade. *Research Policy*. 43, 805-811.
- West, J. and Lakhani, K.R. (2008). Getting clear about communities in open innovation. *Industry and Innovation*. 15 (2), 223-261.

LEARNING PHENOMENA OF MMORPG PLAYERS: A PROPOSED RESEARCH MODEL

Research in Progress

Chollet, Antoine, University of Montpellier, Montpellier, France,
antoine.chollet.pro@gmail.com

Bourdon, Isabelle, University of Montpellier, Montpellier, France,
isabelle.bourdon.pro@gmail.com

Rodhain, Florence, University of Montpellier, Montpellier, France,
florence.rodhain@univ-montp2.fr

Abstract

This research in progress aims to contributing to the already existing knowledge on the phenomena of learning existing in the massively multi-players online role-playing games (MMORPG). The objective is to be able to identify social skills (personality) and professional skills (managerial skills) which the players can develop by playing in MMORPG. A better knowledge of the effects of this technology on the players would allow to know better the consequences on the real life of the players as the recruitment by online games. So, by basing us on learning theories then on a previous qualitative study led with players and old players, we propose a model of research illustrating the influence of the practice of the MMORPG on the player.

Keywords: MMORPG, learning, managerial skills, guild, cognitive constructs of the player.

Introduction

MMORPGs (Massively Multiplayer Online Role Playing Game) are role play games mostly multi-player and were born in the mid-1990s, they were popularized through the game World of Warcraft. The MMORPG market in 2013 weighed almost \$ 15 billion¹, or more than 21% of the worldwide turnover of the video game reaching over \$ 70 billion.

Studies on virtual worlds management are numerous (Berente et al. 2011; Chaturvedi et al. 2011; Goel et al. 2011; Suh et al. 2011) and even encouraged (Wasko et al. 2011). Sometimes spending several hours in front of their screens, MMORPG players do not remain passive and can live unique and sometimes unusual gaming experiences such as virtual funeral for a deceased player in real life². In another words, MMORPGs are increasingly the subject of scientific studies on the development of managerial skills such as leadership (Jang and Ryu 2011; Mysirlaki 2011; Mysirlaki and Paraskeva 2012; Nuangjumnonga and Mitomo 2012; Prax 2014). Experiments in Guild (group of players) in MMORPGs also attract the interest of researchers (Kang et al. 2009; Ang and Zaphiris 2010; Zhong 2011) since they are the crossroads of social interactions and exchanges between players. This is also a true virtual organizations where missions can be leaded under the name instance or raid.

In a previous studies of MMORPG players (Chollet et al. 2013), many testimonies were part of a mobilization phenomenon and learning various skills both cognitive and managerial. Based on this evidence, the study currently aims to provide a research model on the learning phenomenon in MMORPGs. This research's objective requires: (1) identification of specific cognitive factors influencing the player to the learning phenomenon, (2) the definition of managerial skills that can be developed through practice MMORPG, (3) the distinction of moderating variables of the learning phenomenon and (4) the existence of a link between virtual life and real life.

To meet these objectives, we present at first the theoretical framework in which this research is in progress. Then, in a second part, we present the methodology adopted before proposing the research model on which is based the results of the qualitative study.

1 Theoretical Framework

1.1 Interactions between Society and Technology

In view of technological determinism, society does not influence the technique (which evolves alone), but the technique itself can influences society (Ellul 1954; Winner 1977; Zuboff 1988; Vinck 1997; Vendramin and Valenduc 2005). However, many past technologies but also present and future are the result of social interactions but also influences the economic, political and social world. This link between hardware and society refers to the work on the sociomateriality (Orlikowski 2010) considering a construction of the item on the dynamics of social life, "*materiality isn't just something you can resume to things" to the material world. "[. . .] Materiality is imbued with culture, language, imagination, memory. It cannot be reduced to mere objects or pure objectivity"*(Dale 2005, p. 652). This theory is linked to the Social Construction of Technology (SCOT), highlighting the co-construction of nature that occurs between technology and society (Pinch and Bijker 1984).

¹ The Global MMO Games market : payments, intelligence and trends (2013) : <http://www.globalcollect.com/the-global-mmo-games-market>

² Funérailles en ligne pour une joueuse de WoW (2005) : <http://www.journaldugamer.com/2005/11/06/funeraillle-en-ligne-pour-une-joueuse-de-wow/>

1.2 Learning in Society via the Technology

Among the theories of learning, some involve in both learning in society and with technology. Thus, the Social Learning Theory (SLT) (Bandura 1977) emphasizes a learning called "vicarious" which is realized through observation and imitation of physical models, symbolic or imaginary. In addition, through training and group learning, a social facilitation effect is created to improve the performance of the individual. Thus, the SLT emphasizes the importance of the learner's interactions with behavioral, cognitive and environmental factors. Finally, linked to social constructivism and learning theories, the theory Connectivism highlights the contribution of technology in the learning of individuals (Siemens 2005). In other words, the technologies and the current digital age facilitates the linking of individuals and learning, such as learning with a computer only without any physical trainer.

With over 60 years of history (Ichbiah 2009; Cario 2011), the video game technology has evolved so much from a technical point of view (black and white graphics passage to a faithful image synthesis to reality), as a social standpoint (passage of individual game online gambling). Some anecdotes like the Crash of Video Game in 1983 or the transition to 3D with the mid-1990s recall that the video game industry is primarily a flexible technology where innovations but also technological orientation and its use come from both designers as well as players. MMORPGs are thus a contemporary illustration of the previous theories presented above. Our study focuses on the interaction between MMORPG and society via their users in a perspective of social learning. To provide a research model, we conducted an exploratory study and present the different results in terms of literature.

2 Research Methodology

In order to better understand the interaction between technology and MMORPG players, we conducted a qualitative study (Chollet et al. 2013) by semi-structured interview (Wacheux 1996) with 13 players and former players of MMORPG. The topics covered in the interview included: (1) the relationship between the player and the video game technology, (2) the experiences of real-life games and (3) the relationship between the player and his professional life.

The final sample consisted of 9 men and 4 women, aged between 21 to 48 years old. The average age is 31 years old. Out of the 13 individuals, 10 are active and three have stopped MMORPGs. The purpose of this study was to gather the players experiences on MMORPG.

By analyzing the data collected, we conducted a content analysis (Aktouf 1987) in two steps: a thematic analysis and lexical analysis (Fallery and Rodhain 2007). First, a thematic analysis was performed with extraction of verbatim from the corpus of the interviews. A total of 348 categories of verbatim were created and characterized by one or two keywords. A cloud of keywords was carried out to highlight the main themes. To help us, a lexical analysis was performed with the software Alceste. A first analysis with the guild as discriminating variable (if the player belongs or belonged to a guild) has enabled us to bring out four classes with different keywords. The second lexical analysis on the same body but with discriminating variables for individual characteristics, has allowed us to know the profiles assigned to the classes of words. Finally, a reconciliation between the thematic analysis and both lexical analysis was performed (Appendix A). The benefit of this merger is within the confrontation between a thematic manual analysis and automated lexical analysis

3 Proposed Research Model

In order to provide a research model, we rely on previous learning models from literature as well as verbatim excerpts from the qualitative study. Thus, we address first the MMORPG as a learning vector, then the factors of the internal condition in the player before identifying the managerial skills that

can be developed. Finally, we identify moderating factors and influence on the life and work of the player before proposing the research model.

3.1 MMORPG are like Learning Vector in the Player

The study shows first a very ancient complicity between the player and video games, *"it can be said that video games have punctuated my life since I was little"*. For most interviewed players, video games allowed them to grow and have interests and cultural references related to gaming area, *"all my references and my interests [...] gravitate to it and were defined by video games I played when I was a kid"*. The realization of this development with video games and therefore MMORPG illustrate both on the development of the player in his personal life, professional life and also digital or virtual life. This form of learning via technology is reminiscent of the foundations of connectivism, an emerging paradigm in learning theory (Kop and Hill 2008; Bell 2010). Literature already offers several models of research around learning and integrates both the social dimension and the characteristics of the user and the media.

3.1.1 The General Learning Model in Video Games

Inspired by the work of Bandura on social learning (Bandura 1977, 1991), the General Learning Model (GLM) (Appendix B) by Buckley & Anderson (2006) was designed for the media game video and says that learning through the media is a continuous cycle of interaction. First personal variables such as attitudes, beliefs, past experiences, emotions and behavioral tendencies and situational variables characterized by the media itself influences the internal state of the player on three levels: cognition (patterns, attitudes), affect (mood, emotions and behavior) and the psychological awakening of the player (excitement, encouragement, encouragement). The personal and situational variables are grouped under the term "gaming". Finally, the internal condition of the player in turn influences the assessment, decision-making and behavior as a learning phenomenon occurs by the player within the game. These learning situations will end the cycle influencing personal and situational variables. According to the authors: *"People can learn many complex behaviors, attitudes, expectations, beliefs and perception patterns through observation and participation in video games"* (Buckley and Anderson 2006, p. 368).

3.1.2 The Model of Social Meaning in Video Games

Offered by Murphy (2007), the Social Meaning Framework (SMF) is based on the participation of players in online gaming (Appendix C), the MMORPG bringing an evolution into GLM of Buckley & Anderson (2006). According to the author, the GLM does not take into account an important aspect in online games which is the presence of other players and the communication between them. Thus Murphy offers a frame near of GLM where a meaning is created by the interaction between the player, the game and other players.

In developing our research model, we draw from the two existing models in the literature.

3.2 Factors of the Internal Condition in the Player

In accordance with the models presented above, the internal state of the player is an essential step in the learning cycle in MMORPGs. We detail the various concepts identified component inside the player's state.

3.2.1 Pleasure

When a player is playing a MMORPG, one of the main motivations regarding the expected is the pleasure to play or at least the perception that the player has (Chang et al. 2008). Playing with others is

also a primary factor in the pleasure of playing a MMORPGs (Chen et al. 2006). In the exploratory study, the pleasure of playing can both be present: *"I play for the fun of playing"*, seen as a moment of relaxation: *"I play as passionate for fun, to relax. I practice every day [...] is like a trigger"*, but also be lost when the constraints of the game are too strong: *"It may be difficult to imagine, but there is no really fun game, there are only obligations. This to move forever, be the best"*.

3.2.2 Self-Esteem

Several studies on MMORPG introduce the concept of self-esteem (Ryan et al. 2006; Stetina et al. 2011). In our exploratory survey, respondents did not explicitly mentioned self-esteem but the image reflected by the player can be altered: *"in video games, you still have to have a facade because it is hard to be the same in video games and in reality"*, otherwise it can improve the image that the player has of himself: *"I am a little less shy, I discovered that I could be passionate about something and that I could dedicate myself and concentrate over time, something I wasn't so sure about a few years ago"*.

3.2.3 Self-Efficacy

In any video game, the action of a mission or task is an important element in the progress of the game. The feeling of self-efficacy can be mobilized to assess the degree of accomplishment in video games (Lieberman 1997; Klimmt and Hartmann 2006) which in MMORPG (Pavlas 2010). Self-efficacy is also one of the pillars of social learning as defined by Bandura (1994). At several points, the players interviewed have expressed a sense of accomplishment through the MMORPG, both from a personal point of view: *"the online video game, MMO, therefore, taught me to more easily reach out to others [...] be able to be more open to strangers"*, and in terms of challenges to succeed: *"We like to also make challenges that are inside the game: for example as boss killed without killing his henchmen and every time there are rewards. We had made all such deeds in the previous extension"*.

3.2.4 Emotivity

Via the practice of video games, designers incorporate elements allowing the player to experience emotions ranging from fear to joy through pain. The literature on emotions transmitted in video games and felt by the players is quite rich (Boulidoires 2006; Ochs et al. 2009; Geslin 2013). The players interviewed also mention emotions in video games and MMORPG: *"there is more and more emotions, feelings, people become really friends, confidants"* or *"when played much and often [...] inevitably there are emotions that involve them, feelings that apply"* or more generally *"there are games that made me cry, games that have me smile, games that scares me. It's a range of emotions as in daily life"*.

3.2.5 Flow (immersion)

The flow represents the state of immersion reached by an individual in an activity giving it optimum psychological state (Csikszentmihalyi 1991). Several studies on video games consider the flow as an important factor in the study of video game players (Nacke and Lindley 2008; Wang et al. 2008, 2009). To understand if the disposal occurs in the learning phenomenon among the players, and we retain this city concept in the exploratory study, *"it's true that when you are young, we enter more easily into a game and so obviously immersion is easier and you feel the strongest things with more excitement"*.

3.3 Learning Management Skills in the Player

Within the meaning of Bartram (Bartram 2005), managerial skills could be developed in particular by the guild leader (group of players) of their activity into management situations within the meaning of Girin (1990). We will detail the various concepts selected.

3.3.1 Leadership

A previous study (Xanthopoulou and Papagiannidis 2012) already showed a leadership learning phenomenon among MMORPG players being at the head of a guild. The size of the guilds can sometimes go up to several hundred members with a hierarchy. The exploratory study shows several leadership styles adopted by players who have or had positions of responsibility in the game, as a rather selfless style: *"We must learn the strategies on bosses, must be explained, learn to manage a group who will do what..."* or authoritarian: *"We must be very hard when managing people [...] not hard in the unpleasant sense, but in the authoritarian direction and get to impose ourselves"*.

3.3.2 Decision Making

In MMORPGs, players with responsibilities often have to decide on strategies and to make decisions within the moment, as in the instances and raids where the time for reflection is highly restricted. It is therefore important to a guild leader or officer responsible for overseeing the group to make a decision quickly. As already mentioned in the literature (Neto et al. 2011), the decision making in MMORPGs is also clear from the exploratory study, *"with that kind of responsibility, I have learned to have reflexes who, what, where, how have management reflexes systematically in the video game I have"*.

3.3.3 Communication

The community aspect of MMORPG promotes interaction between players and by extension the verbal and written communication. MMORPG thus represent a privileged communication channel for interaction between the players and improves the social aspect in the player (Peterson 2012). The results of the exploratory study show that learning to communicate seem to exist in the MMORPG: *"taught me that communication with other people"*, and take an important place in the game experience, *"in a MMO you have to communicate, it is very important if you want to go on playing"*.

3.3.4 Coordination

In instances and raids, the guild is obliged to coordinate and collaborate to achieve the objective sought (Chen 2009). Each member has a specific role and synergy of everyone's skills allows the success or not of the goal. Coordination enables according Philippette (2014) play well together in MMORPGs. The exploratory study joins those studies like this excerpt: *"The practice of MMO taught mainly to play online and see all the behaviors of players there may be. These are people who are different and these people have a different style of play and I had to learn to coordinate, not necessarily everyone plays the same way but each have the same purpose at the same time"*.

3.4 Moderating Effects on Learning

Several concepts may offset the learning of managerial skills. We will present these moderating effects.

3.4.1 The Role of the Player

In an MMORPG, each player can have one or more roles depending on the features of his character. From experience of the researcher on MMORPG, different roles have been identified (Chollet et al. 2013). The exploratory study also shows the importance of the roles in the game: *"When also manages a guild, [...] it is necessary that everyone plays his role so it is interesting"* It is not uncommon to have several roles, *"I've often had an officer roles, not necessarily important officer roles, but I had a lot of ancillary roles"*.

3.4.2 Addiction

Discussed extensively in literature (Valleur 2006; Kuss and Griffiths 2012; Tisseron et al 2012), addiction in video games may have a moderating role in learning of managerial skills. Some interviewees' players experienced this phenomenon of addiction: "*I was addicted to some online RPGs, I was addicted and it was very hard to pull off*", which can have an impact on the health of the individual "*Some weekends I would not sleep*".

3.4.3 Control Variables

Much research in information systems combine features on users (Jarvenpaa and Staples 2000; Fraser et al. 2001) that could lead to differences in the use of information systems (Gefen and Straub 1997). In the use of MMORPG, we take the characteristics of previous studies (Griffiths et al. 2004; Yee 2006), for example, age, gender, marital status or level of education.

3.5 Influence of MMORPG on the Personal and Professional Life

Video Game Player should not forget the importance of personal, family and professional life. The balance between these can sometimes be a source of disagreement or disturbance as shown in the testimony of one of the interviewees: "*I was scolded by my wife at first because I was leaving, I fled saying "I have a good raid, must I go away" and I was scolded*", or instead as a beneficial link: "*I would love to combine my passion with my mastery of languages, and I'd love to work in the translation of the video game*" or "*if we manage a guild that will teach us necessarily to learn the resources, the team management, it can play on the evolution level because we get to do different things, to be more versatile, and to have some experience in the game that can enable a shift in position*".

These statements join previous studies on balancing work and family life such as that of Ahuja et al. (2007).

3.6 Research Design

After presenting the different concepts, and we propose the following research model for learning managerial skills in MMORPGs next several intrinsic elements to the game and the player (gaming) and extrinsic (internal state of the player, addiction and role) moderated by addiction and role, as well as the link between the player and living real life (Figure 1). This model is based on General Learning Model (Buckley and Anderson 2006) and Social Meaning Framework (Murphy 2007), themselves based on Bandura's Theories (1977, 1986).

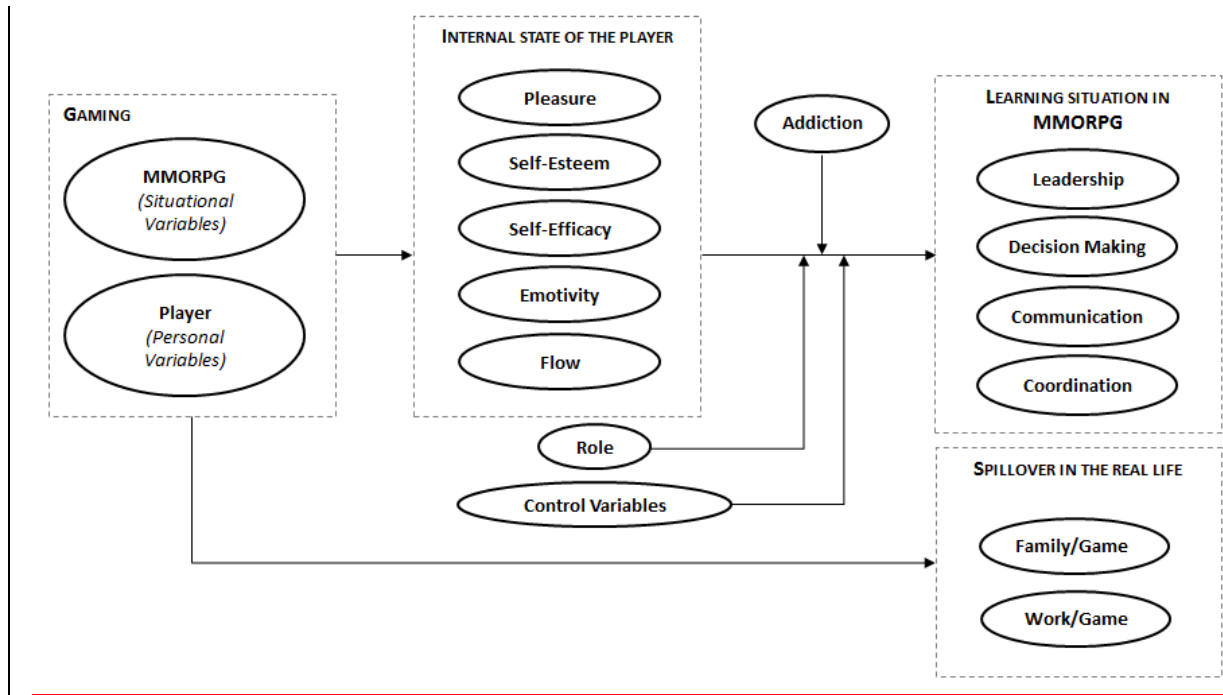


Figure 1. Research Model.

4 Perspectives and Contributions

Already being validated, a quantitative exploratory phase validated the model of different measurement instruments to assess different concepts presented in this research. Data collection for the quantitative phase is over. In total, via an online survey edited thanks to the tool named *Eval and Go*, 707 answers were collected in the exploratory quantitative phase, but only 414 answers are exploitable. For the confirmatory quantitative phase, 3 690 answers were collected but only 2 628 are exploitable. Data processing of the quantitative phase will be able to test hypothesis of the model. For the data processing, the AMOS software will be used. Inspired by the GLM and SMF, this model highlights a learning process of managerial skills by MMORPG players. In 2007, Reeves Professor of Stanford University and a specialist in interactive media, including online games said: "If you want to see what the corporate leadership might look like in three to five years, look whats happening in online games". This statement corroborates an article in the Harvard Business Review³ showing that leadership techniques adopted in video games could step changes on how managers of tomorrow and those of today could improve. A real example of the contribution of MMORPG for the player can be illustrated with the case of Heather Newman, an American of 43 years old, recruited in 2014 in the University of Michigan as the director of communication and information for his talents as a leader in the MMORPG *World of Warcraft* as well as his ability to speak in geek to students thanks to its gaming culture.

³ Leadership's Online Labs (2008) : « <http://hbr.org/2008/05/leaderships-online-labs/ar/1> »

References

- Ahuja, M. K., Chudoba, K. M., Kacmar, C. J., McKnight, D. H., and George, J. F. 2007. "IT Road Warriors: Balancing Work-Family Conflict, Job Autonomy, and Work Overload to Mitigate Turn-over Intentions," *MIS Quarterly* (31:1), pp. 1–17.
- Aktouf, O. 1987. *Méthodologie des sciences sociales et approche qualitative des organisations: une introduction à la démarche classique et une critique*, Sillery: Presses de l'Univ. du Québec.
- Ang, C. S., and Zaphiris, P. 2010. "Social Roles of Players in Mmorpg Guilds," *Information, Communication & Society* (13:4), pp. 592–614.
- Bandura, A. 1977. *Social Learning Theory* (1st edition.), Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. 1986. *Social foundations of thought and action: A social cognitive theory* Prentice-Hall series in social learning theory. (Vol. xiii), Englewood Cliffs, NJ, US: Prentice-Hall, Inc.
- Bandura, A. 1991. "Social cognitive theory of self-regulation," *Organizational Behavior and Human Decision Processes* (50:2), pp. 248–287.
- Bandura, A. 1994. "Self-Efficacy," in *The Corsini Encyclopedia of Psychology*, John Wiley & Sons, Inc.
- Bartram, D. 2005. "The Great Eight competencies: a criterion-centric approach to validation," *The Journal of Applied Psychology* (90:6), pp. 1185–1203.
- Bell, F. 2010. "Connectivism: Its place in theory-informed research and innovation in technology-enabled learning," *The International Review of Research in Open and Distributed Learning* (12:3), pp. 98–118.
- Berente, N., Hansen, S., Pike, J. C., and Bateman, P. J. 2011. "Arguing the Value of Virtual Worlds: Patterns of Discursive Sensemaking of an Innovative Technology," *MIS Q.* (35:3), pp. 685–710.
- Boulidoires, A. 2006. "Le jeu vidéo au regard de la corporéité," *Communication. Information médias théories pratiques* (Vol. 24/2), pp. 141–158.
- Buckley, K. E., and Anderson, C. A. 2006. "A Theoretical Model of the Effects and Consequences of Playing Video Games," in *Playing Video Games - Motives, Responses, and Consequences*, Mahwah, NJ: LEA., pp. 363–378.
- Cario, E. 2011. *Start ! : La grande histoire des jeux vidéo*, Editions de la Martinière.
- Chang, K., Koh, A., Low, B., Onghanseng, D., Tanoto, K., and Thuong, T. 2008. "Why I Love This Online Game: The MMORPG Stickiness Factor," *ICIS 2008 Proceedings*.
- Chaturvedi, A. R., Dolk, D. R., and Drnevich, P. L. 2011. "Design Principles for Virtual Worlds," *MIS Q.* (35:3), pp. 673–684.
- Chen, M. G. 2009. "Communication, Coordination, and Camaraderie in World of Warcraft," *Games and Culture* (4:1), pp. 47–73.
- Chen, V. H.-H., Duh, H. B.-L., Phuah, P. S. K., and Lam, D. Z. Y. 2006. "Enjoyment or Engagement? Role of Social Interaction in Playing Massively Multitplayer Online Role-Playing Games (MMORPGS)," in *Entertainment Computing - ICEC 2006* Lecture Notes in Computer Science, R. Harper, M. Rauterberg, and M. Combetto (eds.), Springer Berlin Heidelberg, pp. 262–267.
- Chollet, A., Bourdon, I., and Rodhain, F. 2013. "Comportements et compétences managériales des joueurs de jeux vidéo en ligne massivement multi-joueurs en situation de groupe : un cas de gamification," in *Actes de la 1ère Journée AIM Serious Game et SI*, Grenoble, p. 30.
- Csikszentmihalyi, M. 1991. *Flow: The Psychology of Optimal Experience*, Harper & Row.
- Dale, K. 2005. "Building a Social Materiality: Spatial and Embodied Politics in Organizational Control," *Organization* (12:5), pp. 649–678.
- Ellul, J. 1954. *La technique, ou, l'enjeu du siècle* (1ere edition.), Paris: Armand Colin.
- Fallery, B., and Rodhain, F. 2007. "Quatre approches pour l'analyse de données textuelles: lexicale, linguistique, cognitive, thématique," in *16ème Conférence Internationale de Management Stratégique*.

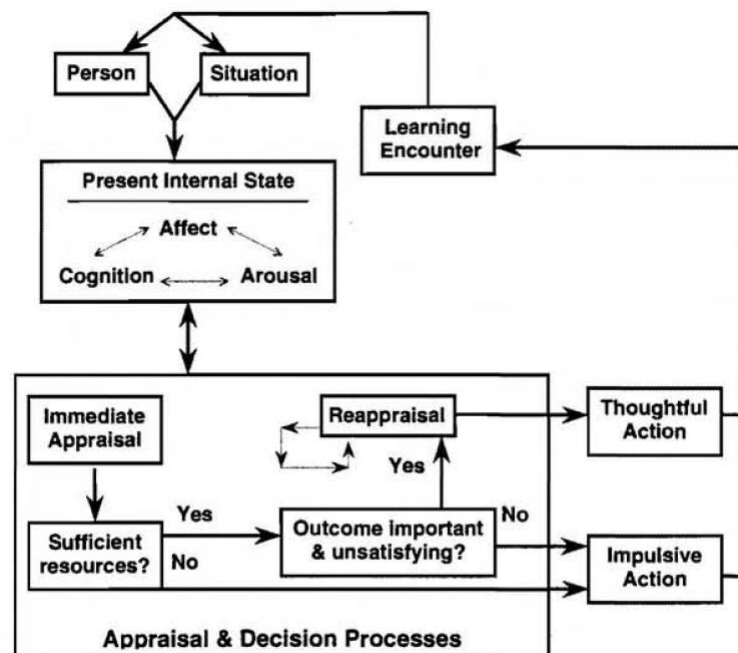
- Fraser, V., Marcella, R., and Middleton, I. 2001. "Employee perceptions of knowledge sharing: Employment threat or synergy for the greater good? A case study," *Competitive Intelligence Review* (11:2), pp. 39–52.
- Gefen, D., and Straub, D. W. 1997. "Gender Differences in the Perception and Use of E-Mail: An Extension to the Technology Acceptance Model," *MIS Quarterly* (21:4), pp. 389–400.
- Geslin, E. 2013. "Processus d'induction d'émotions dans les environnements virtuels et le jeu vidéo," Ecole nationale supérieure d'arts et métiers - ENSAM.
- Girin, J. 1990. "L'analyse empirique des situations de gestion: éléments de théorie et de méthode," *Epistémologies et sciences de gestion, Economica*, pp. 141–182.
- Goel, L., Johnson, N. A., Junglas, I., and Ives, B. 2011. "From Space to Place: Predicting Users' Intentions to Return to Virtual Worlds," *MIS Q.* (35:3), pp. 749–772.
- Griffiths, M. D., Davies, M. N. O., and Chappell, D. 2004. "Online computer gaming: a comparison of adolescent and adult gamers," *Journal of Adolescence Video Games and Public Health* (27:1), pp. 87–96.
- Ichbiah, D. 2009. *La Saga des Jeux Vidéos* (5e Revue et augmentée.), Pix'N Love Editions.
- Jang, Y., and Ryu, S. 2011. "Exploring Game Experiences and Game Leadership in Massively Multiplayer Online Role-Playing Games: Exploring Game Leaderships in MMORPGs," *British Journal of Educational Technology* (42:4), pp. 616–623.
- Jarvenpaa, S. L., and Staples, D. S. 2000. "The use of collaborative electronic media for information sharing: an exploratory study of determinants," *The Journal of Strategic Information Systems* (9:2–3), pp. 129–154.
- Kang, J., Ko, I., and Ko, Y. 2009. "The Impact of Social Support of Guild Members and Psychological Factors on Flow and Game Loyalty in MMORPG," in *42nd Hawaii International Conference on System Sciences, 2009. HICSS '09*, Presented at the 42nd Hawaii International Conference on System Sciences, 2009. HICSS '09, , January, pp. 1–9.
- Klimmt, C., and Hartmann, T. 2006. "Effectance, Self-Efficacy, and the Motivation to Play Video Games," in *Playing video games: Motives, responses, and consequences* P. Vorderer and J. Bryant (eds.), Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, pp. 133–145.
- Kop, R., and Hill, A. 2008. "Connectivism: Learning theory of the future or vestige of the past?," *The International Review of Research in Open and Distributed Learning* (9:3).
- Lieberman, D. A. 1997. "Interactive video games for health promotion: Effects on knowledge, self-efficacy, social support, and health," in *Health promotion and interactive technology: Theoretical applications and future directions* LEA's communication series., interactive video games for health promotion: Effects on knowledge, self-efficacy, social support, R. L. Street, Jr., W. R. Gold, and T. R. Manning (eds.), Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, pp. 103–120.
- Murphy, S. 2007. "A Social Meaning Framework for Research on Participation in Social Online Games," *Journal of Media Psychology* (12:3), pp. 13–23.
- Mysirlaki, S. 2011. *Leadership in a Networked World: The Case of Massive Multiplayer Online Environments*.
- Mysirlaki, S., and Paraskeva, F. 2012. "Leadership in MMOGs: A Field of Research on Virtual Teams," *Electronic Journal of e-Learning* (10:2), pp. 223–234.
- Nacke, L., and Lindley, C. A. 2008. "Flow and Immersion in First-person Shooters: Measuring the Player's Gameplay Experience," in *Proceedings of the 2008 Conference on Future Play: Research, Play, Share* Future Play '08, New York, NY, USA: ACM, pp. 81–88.
- Neto, H. C. S., Carvalho, L. F. B. S., Paraguaçu, F., and Lopes, R. V. V. 2011. "A MMORPG Decision-Making Model Based on Persuasive Reciprocity," in *Integrated Uncertainty in Knowledge Modelling and Decision Making* Lecture Notes in Computer Science, Y. Tang, V.-N. Huynh, and J. Lawry (eds.), Springer Berlin Heidelberg, pp. 33–47.
- Nuangjumnonga, T., and Mitomo, H. 2012. "Leadership development through online gaming,"
- Ochs, M., Sabouret, N., and Corruble, V. 2009. "Simulation de la dynamique des émotions et des relations sociales de personnages virtuels," *Revue d'intelligence artificielle* (23:2-3), pp. 327–357.

- Orlikowski, W. J. 2010. "The sociomateriality of organisational life: considering technology in management research," *Cambridge Journal of Economics* (34:1), pp. 125–141.
- Pavlas, D. 2010. "A model of flow and play in game-based learning: The impact of game characteristics, player traits, and player states," University of Central Florida, Orlando, Florida.
- Peterson, M. 2012. "Learner interaction in a massively multiplayer online role playing game (MMORPG): A sociocultural discourse analysis," *ReCALL* (24:03), pp. 361–380.
- Philippette, T. 2014. "Bien jouer ensemble : les activités de coordination des joueurs de jeux de rôle en ligne massivement multijoueurs," Faculté des sciences économiques, sociales, politiques et de communication.
- Pinch, T. J., and Bijker, W. E. 1984. "The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other," *Social Studies of Science* (14:3), pp. 399–441.
- Prax, P. 2014. "Leadership Style in World of Warcraft Raid Guilds,"
- Ryan, R. M., Rigby, C. S., and Przybylski, A. 2006. "The Motivational Pull of Video Games: A Self-Determination Theory Approach," *Motivation and Emotion* (30:4), pp. 344–360.
- Stetina, B. U., Kothgassner, O. D., Lehenbauer, M., and Kryspin-Exner, I. 2011. "Beyond the fascination of online-games: Probing addictive behavior and depression in the world of online-gaming," *Computers in Human Behavior* Current Research Topics in Cognitive Load Theory Third International Cognitive Load Theory Conference (27:1), pp. 473–479.
- Suh, K.-S., Kim, H., and Suh, E. K. 2011. "What if Your Avatar Looks Like You? Dual Congruity Perspectives for Avatar Use," *MIS Q.* (35:3), pp. 711–730.
- Vendramin, P., and Valenduc, G. 2005. "Technologie et société, destins croisés,"
- Vinck, D. 1997. *Sociologie des Sciences*, Paris: Armand Colin.
- Wacheux, F. 1996. *Méthodes qualitatives et recherche en gestion*, Economica.
- Wang, C. K. J., Khoo, A., Liu, W. C., and Divaharan, S. 2008. "Passion and intrinsic motivation in digital gaming," *Cyberpsychology & behavior: the impact of the Internet, multimedia and virtual reality on behavior and society* (11:1), pp. 39–45.
- Wang, C. K. J., Liu, W. C., and Khoo, A. 2009. "The Psychometric Properties of Dispositional Flow Scale-2 in Internet Gaming," *Current Psychology* (28:3), pp. 194–201.
- Wasko, M., Teigland, R., Leidner, D., and Jarvenpaa, S. 2011. "Stepping into the Internet: New Ventures in Virtual Worlds," *MIS Q.* (35:3), pp. 645–652.
- Winner, L. 1977. *Autonomous Technology: Technics-out-of-control as a Theme in Political Thought*, MIT Press.
- Xanthopoulou, D., and Papagiannidis, S. 2012. "Play online, work better? Examining the spillover of active learning and transformational leadership," *Technological Forecasting and Social Change* (79:7), pp. 1328–1339.
- Yee, N. 2006. "The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments," *Presence* (15:3), pp. 309–329.
- Zhong, Z.-J. 2011. "The effects of collective MMORPG (Massively Multiplayer Online Role-Playing Games) play on gamers' online and offline social capital," *Computers in Human Behavior* (27:6), pp. 2352–2363.
- Zuboff, S. 1988. *In the age of the smart machine: The future of work and power* (Basic Books.), Basic Books.

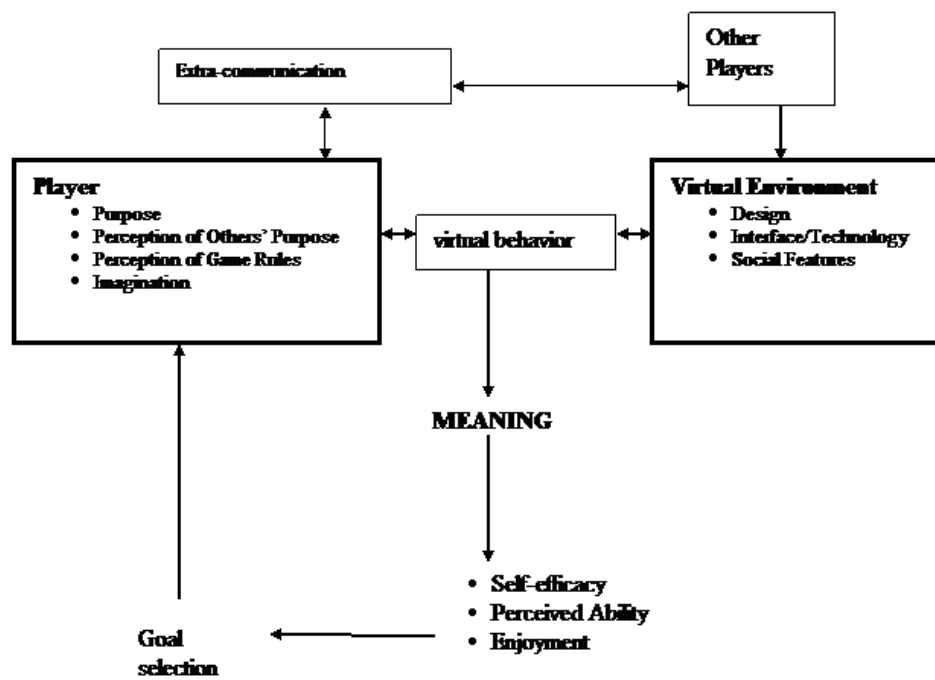
Appendix A: Correlation of Thematic and Lexical Analysis

| ANALYSE 1 <i>Thématique</i> | ANALYSE 2-A <i>Lexical with "guild" variable</i> | ANALYSE 2-B <i>Lexical with "personal" variables</i> | SYNTHESIS |
|--|--|--|-------------------------------|
| Main keywords | 5 keywords corresponding to the class | Variables discriminating classes corresponding | Main thematic obtained |
| Family, Everyday life, Professional life, Personal life, Love, Social, Personality, Friendship, Relationship, Generation, Human, Violence, Environment, Blooming | Classe 3 : People, Generation, Young person, English, environment | Classe 3 : Woman, Master of Arts or Science, Phd, Active Players Classe 4 : A Levels, 31-60 years, Inactive Players | Social |
| Art, Culture, Language, Emotion, Game, Leisure | Classe 2 : Art, Purpose, Concept, Independent, Cinema | <i>Aucune classe correspondante</i> | Art |
| Technology, Practice, Money, Activities, Industry | Classe 1 : PC, Month, World of Warcraft, Play, to Buy | Classe 1 : BTEC First Diploma, BTEC HND, 20-30 years | Technology |
| Conflict, Guild, Competition, Learning, Time, Experience, Change, Skill, Communication, Behavior, Community | Classe 4 : Work, MMO, Colleague, Job, Learn | Classe 2 : Man, Active Players, 20-30 years | Management |

Appendix B: General Learning Model (Buckley and Anderson, 2006)



Appendix C: Social Meaning Framework (Murphy, 2007)



UNDERSTANDING KNOWLEDGE SHARING IN HEALTH CARE SYSTEM

Complete Research

Ben Hamouda, Iman, Telecom Business School, France, iman.ben_hamouda@telecom-em.eu

Feki, Mondher, Telecom Business School, France, mondher.feki@telecom-em.eu

Boughzala, Imed, Telecom Business School, France, imed.boughzala@telecom-em.eu

Chourabi, Olfa, Telecom Business School, France, olfa.chourabi@telecom-em.eu

Abstract

To date, healthcare organizations need well-established knowledge to support care quality. Managers invest in motivating medical staff and associated healthcare stakeholders to share their respective knowledge. Additionally, they invest in advanced technology to facilitate sharing process. However, these efforts essentially raise classification and privacy concerns. Technology configuration that does not meet management process requirements may lead to managerial difficulties. In fact, information classification in healthcare systems should respect privacy, yet still be accessible. Using qualitative research, our paper proposes a novel ontological model that extends the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) with privacy dimension to enhance access to sensitive patient's data.

Keywords: Knowledge Sharing, Classification, Health Care, Ontology.

1 Introduction

Knowledge sharing plays a determining role in improving organizational performance. Indeed, generating knowledge, and sharing it, provides information that supports care quality and decision-making. It can also engender new ideas, which create business value. Organizations increasingly invest in knowledge sharing systems to capture, circulate, and reuse information resources effectively and efficiently (Nonaka and Takeuchi, 1995; Davenport and Prusak, 1998; Teigland, 2003). This process is very important in the healthcare sector where exchanges of information and knowledge are critical to patient management.

In the health sector, there are wide ranges of data relative to patients, diseases, treatments, analysis results, etc. These data, which occurs quickly in massive quantities, represent a big data. This term describes “large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information.” (Gandomi and Haider, 2015, p.138).

Big data in healthcare can play a crucial role with epidemics’ prediction, curing diseases, health monitoring, Pharmacovigilance, personalized medicine, and an overall improved quality of life. However, it also poses threats to a patient’s privacy, and raises many new privacy concerns. This creates a challenge to make a balance between knowledge and information sharing, allowing for value creation and protection of patient data.

Even though knowledge sharing’s impact on decision making and improvement to healthcare services has been discussed in previous research, there are some controversial factors that reduce tendencies to share within healthcare systems. Attitude, culture, privacy concerns, and technologies have all been identified as barriers to the knowledge sharing process (Choi et al., 2008; Witherspoon et al., 2013).

To protect healthcare information and facilitate knowledge sharing processes in this sector, it is critical to define convenient classification principles. However, in France, medical data and information do not have different classifications; they are all confidential and protected by medical secrecy. The Touraine Act (Article 47 of the Health Bill) would enable researchers, associations, and private companies access to billions of types of medical data. Although, classification rules create distinct classes of information based on their levels of sensitivity, as well providing access to qualified information demand a strategic value.

This paper’s goal is to provide a data classification framework that enables knowledge sharing within healthcare systems. Our research question is the following: How can we improve data sharing in the healthcare field within big data context?

This paper is organized as follows. We begin by presenting the various present researches about information and knowledge sharing, and then we move on to the current status of big data and the current context of the healthcare sector. Afterwards, we describe the research methodology used to answer our question followed by the presentation of the results. Then we discuss the main findings and propose a structural ontology. We conclude by underlining the limits of our research.

2 Background

Knowledge sharing (KS) is defined as “*the act of making knowledge available to others within the organization. Knowledge sharing between individuals is the process by which knowledge held by an individual is converted into a form that can be understood, absorbed, and used by other individuals*” (Ipe, 2003, p. 341). It is an exchanging process of implicit and explicit knowledge allowing the development of new ideas and subsequently the improvement of decision-making and organizational performance (Nonaka and Takeuchi, 1995; Cummings, 2004).

Knowledge is essentially related to human action (Nonaka and Takeuchi, 1995). Furthermore, it can be distinguished between the tacit and the explicit dimension of knowledge, which presents a way of

articulating knowledge (Baumard, 1999). Recognition of the differences between tacit and explicit knowledge is essential when considering the transferability and knowledge sharing (Boughzala, 2007). The passage of an explicit knowledge into a tacit knowledge is translated by information sharing.

Various factors affect decisions about knowledge sharing among employees and other organizational participants; however, previous research divides these factors into two categories:

- Most previous studies emphasize the role of attitude and behaviors on knowledge sharing because they depend heavily on human choices and tendencies. They have described knowledge sharability, knowledge sharing, and how they correlate within an intra organizational setting (Bock et al., 2005, Wang and Noe, 2010, Chennamaneni et al., 2012, Boughzala and Briggs, 2012). Individual perceptions of knowledge shareability (Cabrera and Cabrera, 2005) have been shown to correlate with a number of factors such as intrinsic and extrinsic motivation (Osterloh and Frey, 2000); awareness of conflicts of interest or vulnerability (Argote et al., 2003); physical capability to share, and certain personality traits including: self-interest and investment and a personal disposition to sharing knowledge (Matzler et al., 2011).
- The ability of information technology to facilitate knowledge sharing processes from storing to generating knowledge is critical as well. Technology empowers the individual, and significantly impresses upon them the attitude of knowledge sharing (Hinds et al., 2003). Cabrera and Cabrera (2005) stress well-designed technologies and user-friendly applications as important factors that simplify tasks, reduce completion time, and the perceptions of cost. Jarvenpaa and Staples (1999) discuss that “*ergonomic technologies influence people’s tendencies to engage in knowledge sharing within an electronic information exchange*”. Sophisticated tools like groupware, backup databases, communicating networks, knowledge management systems, workflow technology, and more-recently social network-based systems, support the knowledge exchange process (Ruggles, 1998; Kankanhalli et al., 2005; Hung et al., 2013). Although information technology has a considerable impact on developing of knowledge sharing, in many instances presenting new technology has failed because of inadequate attention to existing organizational cultures, insufficient adoption and untrained personnel (Cabrera and Cabrera, 2005).

Organizations face up to massive data flow, which contains potential value. The sharing of these data coming from various sources creates the big data, often characterized by five Vs: volume, variety, velocity, veracity and value. Fosso Wamba et al. (2015, p. 2) defines big data as “*a holistic approach to manage, process and analyze 5 Vs in order to create actionable insights for sustained value delivery, measuring performance and establishing competitive advantages.*” Wang et al. (2015) study the impact of big data on enhancing healthcare services. They find that big data technology in healthcare should contain these attributes: traceability of data, analyzing unstructured data, and qualified analysis for patterns of treatment that support decision-making.

The organizations, which accumulate big data, are going to have to display solutions to manage, analyze and interpret these huge volumes of data but also make sure that these data are preserved in complete safety. Indeed, the privacy issues of medical data worry the patients. The patients do not want that their data to be shared (Perera et al., 2011). However, the sharing of these data is useful. For example, Davis et al. (2014) showed the benefits of sharing psychological electronic records for the care of cancer. The question that arises in this context is “what type of data to share and with whom” (Borgman, 2015). Concerns about privacy are divided into: the type of information stored (e.g., medical, biometric, financial, behavioral, and biographical) and sector would use the shared information (Smith et al., 2011). Other researches have insisted on the weakness of used technology in health sector to handle the issues of privacy (Cabrera et al., 2005; Witherspoon et al., 2013).

Legally, there is no typology permitting for the operation and classification of medical data. However, in the absence of a clear legal list and in the absence of typology of the data, the above, evoked definitions, are only extrapolations. Thus there is a need for text that is susceptible to developing a clear definition that outlines each type of data. By doing this, it would be easier for each party to know the

nature and the significance of the information they have. It could particularly be utilized to create a distinction between the various data at the medical level. This would also permit the application of a different system of access. Certain information, considered more sensitive than others, would not automatically be communicated to certain members of the medical team.

This typology would be conducted more and would then provide a balance between sharable information and privacy issues. The persistence and quality of care would be assured without spoiling the rights of a patient, and continuing to respect their private lives. Ultimately, this would retain the privacy of the information in medical materials.

3 Methodology

According to our literature review, we notice that the concept knowledge sharing and privacy were rarely studied jointly and especially in the healthcare sector. Given that few studies have handled these two themes in an explicit way, a qualitative exploratory research turns out to be useful in this context in order to discover the characteristics and the articulation between these concepts.

For this purpose, we adopted an exploratory approach. It is an inductive approach that consists of exploring a phenomenon by rejecting all the previous knowledge (Thietart et al., 2014). In this work, our purpose is to build a structural ontology.

We led non-directive, face-to-face interviews with physicians working in different hospitals. The guide of interview consists of 3 themes: (1) Knowledge sharing and value, (2) Problems relative to medical information sharing, and (3) Privacy - Classification. It was tested and refined with a physician, a department head of resuscitation and intensive care, who participated in several projects involving the implementation of the Computerized Patient Record.

The appeal of a non-directive interview, answers two essential objectives of the qualitative inquiries: depth and objectivity (Frisch, 1999). This mode of data collection is useful when the object of research is less defined or poorly structured, which is our case (Thietart et al., 2014).

The data analysis was performed by using the qualitative content analysis method (Berelson, 1952) supported by NVivo 10 computer software.

Our sample consists of 4 physicians who were requested to participate in our study from our personal networks. These physicians practice in 4 different hospitals in France. Their experience varies between 20 and 30 years. Their profiles vary from general practitioner, anesthetist - resuscitator, specialist in vascular medicine, and internal medicine. These physicians have participated in implementing of medical record in many hospitals in France.

After the transcription of fourth interviews, we made a content analysis supported by Nvivo, which gave rise to 372 passages of text. These passages were listed according to our four themes.

The coding was separately made by three authors by using the Nvivo software. The purpose was to compare and discuss the results obtained by each until they obtained a consensus. The results were then validated by the senior researcher and translated into English by two native English speakers.

4 Results

This section presents an analysis of the transcriptions in the light of themes.

4.1 Knowledge sharing and value

According to the interviewed physicians, sharing data, information, and medical knowledge in a hospital setting is critical. This sharing of information allows for improvement in the precision of a diagnosis, and care quality.

The first interviewed physician stated the criticality of sharing the medical history of a patient in order to determine whether or not to resuscitate them; *“Let us imagine that I have a patient who is a hemo-*

philiac and that I have no information and that we operate on him without knowing this. It is crucial that this information is disclosed; otherwise the results could be disastrous. The sharing of information is compulsory. We cannot take the necessary decisions if we do not share medical information and knowledge”.

The second physician has a similar opinion; *“Yes, we can say it all the same that sharing improves the activity of care”.*

The third physician asserted that there is a relation of reciprocity between sharing and quality; *“sharing and improvement are mutually important. One does not go without the other. Naturally, if we want to improve the quality of the care, it is necessary to share”.*

For the fourth physician, it is important to permanently share the information and knowledge among a team in order to improve collaboration. The objective of a large amount of information sharing is to improve the quality of the care. *“It is tightly in line with the aim of improving the quality of treatment, the quality of an action, which conveys for me the necessity of having the maximum amount of information, in order to make experienced decisions”.*

4.2 Problems relative to medical information sharing

In this part, we present the problems relative to the information sharing at the levels of technology and sharing context.

4.2.1 Technology

Four interviewed physicians explained the obstacles that can surround sharing between medical staff. They expressed the problems bound to a computerized patient file. This tool, in theory, is considered to favor information sharing within hospital organizations. However, according to these physicians, the ignorance and unwillingness to cooperate interrupts the implementation of sharing technology and the incentives to integrate within a collaborative work.

For the first physician, the problem is that the hospital executives bring a very general tool and it does not meet the expectations of professionals; *“it is a big problem because often management provides a very general IT tool but after personalization problems appear for us. And if we do not personalize it, we do not use it properly”.* He noted that the information technology program is not successful or optimal enough.

The second one emphasized the same problems with the computerized patient record; *“this computerization of the patient record raises problems rather than solving the previous difficulties.”*

For the third physician, following the dematerialization of medical records, he noticed that the practices of the physician staff were standardized and consequently the contents of the patient file were reduced. These files were replaced with simple, basic documents that often limit detail. The computerized patient file, according to him, is not made to make the necessary reflections about a disease or the quality of the treatment of a patient; *“It is true that information is a sort of mechanical data of the individual at the first moment, which makes the patient a summarized concept of biomedical dimensions. Then we share this information by use of computerized patient record. To sum it up, all of the mechanisms simply provide biomedical dimensions”.*

The fourth physician noticed that the medical staff is not trained to use the computerized patient file. Thus, they resist this use and this type of sharing. In addition, there are certain medical procedures, which cannot use the software and IT applications; they were not intended to be used for certain medical acts. Some information stays on a paper file and thus the medical data is divided between paper and computerized files. *“The software won’t be used as it’s not always the most practical. Certain data is missing as we do not manage to include it in the system ... Thus the tool is not adapted enough and the staff is not trained.”* He suggests making the software more useful for medical activities by integrating data and statistical studies which can improve the physician’s knowledge.

4.2.2 Sharing context

The hospital organization is characterized by very complex activities and is subjected to strong pressures to integrate technological innovations. The interviewed physicians cited some characteristics of this organization, which can occasionally limit sharing, and asked them about the notion of security.

The first physician explained why some medical staffs do not use the computerized patient file and thus do not receive more pertinent and related information. He expressed that this is useful for care giving, *“In the hospital, there are a lot of medical staff who pass-through and sometimes just for temporary missions. So they do not have time to learn how to use software.”* In regards to the security of the data and the information, in theory, the IT organization in terms, depends on every critical structure to put barriers and protect the medical data. However, in practice, it is not easy to set up a system, which allows us to ensure the control of the data; *“actually at the moment it is not very easy to put up protection barriers for data security. How to implement them? According to the profile of each one? That can be made, completely. Because it almost impossible to control medical staff participants, who comes a long time or for a temporary moment, and setting privacy rules is complicated”*.

The second physician explained that, in hospitals, security systems are not standardized. Theoretically, the information is not accessible to everyone, but in practice, it is completely the opposite; *“Even when you have security systems with access controls, generally the passwords are hung on the wall, like everywhere. If you want the hospital is the place of ‘3*8’ (shifting work time), there is always a team which comes after the previous one, it is a place where there are temporary employees, it is a place in which the physicians and the interns change all of the time.”*

The third physician works in two organizations, a private clinic and a hospital. At the hospital, he had no training on the use of medical software and he is a just a practitioner for consultations; *“I can’t find time to learn the technology or other irrelevant points. Because it seems very complicated to me, and so I continue to use paper.”* He also indicated the problem associated with the security of data in the hospital; *“as the files are computerized, then all the reports would be in computer programs. A physician in another department always has the leisure to read the diverse and varied reports on a patient even if that does not concern him”*.

The fourth physician indicated that the communication of medical data can sometimes, damage the private life of some patients, especially if this data is shared with insurances or banks. In regards to the communication of confidential data, the physician sometimes tries to hide the identity of the patient: *“When the patient does not want me to communicate their information, it isn’t a fault that I talk with my colleagues without giving their name, it is true that it remains anonymous but it is a partial anonymity...”*

4.3 Privacy - Classification

In regards to the security of confidential data, the four physicians gave very contradictory answers; however, at the same time they consider that medical information is quite confidential. Much like the law has considered it for a long time, and they believe that they have to respect the patient will to avoid their information being broadcasted. However, at the same time, for them, to assure the quality of care, all the information must be shareable between the medical staff. Sometimes for certain legal reasons, information is not marked on the file but rather shared verbally.

According to the first physician, certain data must not be shared or distributed within the organization. He noticed that the IT tool does not propose solutions to protect and classify this type of data; *“when we want to keep information a secret, sometimes we use numbers instead of the names of the patients... This type of information is called confidential. This should be defined as distinct information, but with whom can we share it? Which information is shareable? This would be fine if the technology could state it.”*

He explained also the complexity of sharing sensitive data; *“I remember when I was in ‘La Reunion’, people were affected by HIV and accessing their file was impossible. This point made a complicated*

situation. For these types of patients, stricter accessibility rules have to be defined for other physicians, moreover, which physicians can access the file should be based off of the patient's agreement. Therefore, more recently we share sensitive information verbally. Because, at the moment this is still not computerized very well."

For the second physician, the decision about classification of medical data returns to the patient, the patient is the owner of his data and only he can define how his data can be shared between the medical staff. "It is not me, as his physician, it is the patient. Me, I don't care, it is the patient who must choose what is available to the general public and what is not."

The third physician defined the information, which is not shareable and the people who can share them: "So, there are confidential things which cannot be shared by everybody. This information relates to personal patient life, for instance, mental elements...because we ask a patient about choosing which information is confidential and which is private. Sometimes we are involved in a huge flood of information, and don't know how much of it should be shared." In regards to the confidential data of the patient, he does not note them on the computerized file and if necessary he shares it verbally with his colleagues. According to him "The professionals share their knowledge. However, there are questions about which knowledge and information should be shared and in which way it should be done?"

The fourth physician explained that medical staffs may have access to very private and highly confidential information, and should respect professional secrecy. Additionally, he mentioned that some patients may require certain private or confidential information to be omitted from a file. He begins, "I don't want to note the name of disease at all, because it is so private. And then it can be a demand. Some patients ask me not to note it on their file. This can be, I'll give you some examples, a voluntary interruption of a pregnancy, or this can be sexual assault. Of course it is medical issue, however, it is a personal concern and we can't share it". He noted that the patient file sometimes contains information that is not useful or relevant to ensuring the consistency of care, and it is marked by his colleagues.

5 Discussion

The literature review shows the role of knowledge sharing in the improvement of care quality and value creation. Interviewed physicians stressed the problems associated with adapting to a new technology for knowledge sharing, which slows the process down or in some cases, is totally ignored by the medical staff. The lack of skills and knowledge to operate a technology create a dilemma. The medical staff was not trained to use information technologies. Interviewed physicians noticed that the computerized patient file is not adequate to the characteristics of their activities. They underlined that the technology is only a storage place; the studies statistics, the reporting or other elements to improve upon the reflection, are not integrated into this technology.

The other difficulty involving the use of technology is about the massive amount of data recorded which is generally unrelated, confusing, and medical staff have a problem finding meaningful knowledge from computerized records. The participants, in our research, emphasized that without the personalization of shared information in databases, it is not possible to take advantage of recorded files and it just wastes the physicians' time. Expanding on this, finding useful knowledge is another dilemma. The physicians stated that although electronic records shared with them can provide access to various information, without converting it to qualified knowledge and intelligence, technologies do not improve treatment services. In this case, technology causes a reduction in the knowledge sharing attitude.

The security issue is the one of the main concerns that was discussed by participants. Patients have various concerns about their personal data and are sometimes reluctant to share their information. Theoretically, the information is not accessible to everyone, but in practice, it is the complete opposite. Access to these files requires a password, and frequently these passwords are hung on the wall and everyone can access these databases. The professional secrecy principal is not able to cover all privacy

issues in the healthcare system, because in certain cases there is no distinct rule to describe which information is shareable and which is not.

Participants substantially discussed determining which information is shareable. They state that the knowledge and information sharing improve decision-making processes. However, security concerns represent a predicament in determining what should be shared and what should be private. The patients determine decisions about which data are shareable. Currently there are no principles to classify information as confidential, or sensitive, this causes physicians to block access and consequently blocks sharing too. Furthermore, information technology does not propose qualified solutions for considering both privacy issues and shareability processes.

Ontologies represent one technology that can contribute to this goal. Ontologies allow healthcare knowledge sharing through a semantic access to computerized patient record. e.g. by defining clinical terminologies for precise and sharable expressions in patient record entries. Another advantage of ontologies is that their hierarchical structure will result in better control over access and use of personalized medical information addressing privacy issues.

Data come into the big data engine from an increasingly wide variety of sources. The data can be confidential or publicly available; it can be personal data relating to individuals; or being the combination of various issues. The classification is the intermediate step between the big data platform and the business analytics. A combination of the classification principles with big data functions can respond to the patients' security and medical staff concerns at the same time.

To sum up this exploratory study, we propose a novel ontological structure for electronic patient record annotation.

6 Proposal of a structural ontology: OntClassHealth

As the use of computerized records increases, healthcare stakeholders should be expected to make the best possible use of their information.

Our research results reveal that the current structure of the computerized patient records does not allow effective and secure knowledge sharing. This record has been built to solve sharing issues and to improve collaborative work. Yet, we noticed that there is a gap between expectations regarding a computerized patient records and its current use.

Health care domain is gaining more focus in the field of Ontology and Semantic research (Yuwen et al., 2010; Prathima et al., 2011). Ontology provides unique representation of knowledge. It enables unambiguous recording of data in a knowledge base. Medical Ontologies such as Medical Language System (UMLS) (Brut et al., 2011) and SNOMED would improve the access to the electronic patient records privacy.

However, the privacy dimension is not covered in the existing medical ontologies. In this paper, we propose a novel ontological structure that extends the existing medical ontologies with computerized patient record structure with privacy dimension.

Computerized patient records could be accessed through a dual semantic annotation based on medical ontology for representing patient diseases and on computerized patient records structure for representing the information privacy level. The proposed privacy classification levels were defined based on our interviews. Figure 3 (Appendix), presents this ontology called "OntClassHealth". We have adopted a design-based methodology (Isaac, 2005), for building ontology concept hierarchies. Each item of the computerized patient record is annotated as "public", "private" or "confidential" based on the patient record structure (Figure 1) and SNOMED (Figure 2) domain concepts. An annotation example is presented in Figure 3.

7 Conclusion

This study approaches the role of the classification on the knowledge sharing in the healthcare sector. The finding of this qualitative research shows that the decision of classification must be shared with the patient to improve the information sharing. However, one of limitations of this study is that the size of the sample is small. The exploitation of collected data is difficult to act upon, and we cannot generalize our results. Notwithstanding the above, this study proposes ontology based electronic patient record annotation, as it relates to medical data. Experimental research focusing on our proposition would shed further light on its scientific utility.

References

- Argote, L., B. McEvily., R. Reagans (2003). "Managing Knowledge in Organizations: An Integrative Framework and Review of Emerging Themes". *Management Science*, vol 49; 4, 571-582.
- Baumard, P. (1999). Tacit Knowledge in Organizations. *Business & Economics*, vol 264.
- Berelson, B. (1952). "Content Analysis in Communication Research", New York, University Press.
- Bock, G. W., R. W. Zmud., Y. G. Kim., J. N. Lee (2005). "Behavioral Intention Formation in Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social- Psychological Forces, and Organizational Climate". *MIS Quarterly*, vol 29:1, 87-111.
- Boughzala, I. (2007). *Ingénierie de la collaboration, Théorie, technologies et pratiques. Collection technique et scientifique des télécommunications*, Hermes-Lavoisier, Paris.
- Boughzala I. and R. O. Briggs (2012). "A value frequency model of knowledge sharing: an exploratory study on knowledge sharability in cross-organizational collaboration". *Electronic Markets*, vol 22:1, 9-19.
- Borgman, C. (2015). "Why are data sharing and reuse so difficult to do?", Stanford Data Day , Los Angeles.
- Brut, M., D. Al-Kukhun., A. Péninou., M-F. Canut., F. Sèdes (2011). "Structuration et Accès au Dossier Patient Médical Personnel: approche par ontologie et politique d'accès XACML".
- Cabrera, E. F. and A. Cabrera (2005). Fostering knowledge sharing through people management practices. *The International Journal of Human Resource Management*, vol.16, issue.
- Chennamaneni, A., T.C.J. Teng., M.K. Raja (2012). "Unified model of knowledge sharing behaviours: theoretical development and empirical test". *Behaviour & Information Technology*, Vol. 31, No. 11, 1097–1115.
- Choi, S.Y., Y. S. Kang., H. Lee (2008). "The effects of socio-technical enablers on knowledge sharing: an exploratory examination". *Journal of Information Science*, vol 34:5,742-754.
- Cummings, J. N. (2004). "Work groups, structural diversity, and knowledge sharing in a global organization". *Management Science*, vol 50:3, 352-36.
- Davis, E., B. Thewes., P. Butow., C. Mason (2014). "A qualitative study of consumer attitudes to sharing psychosocial information within the multidisciplinary cancer care team". *Support Care Cancer*, vol 22(4), 869–70.
- Davenport, T. H., and L. Prusak (1998). "Working Knowledge: How Organizations Manage What They Know." *Harvard Business School Press, Boston*.
- Frisch, F., 1999. *Les Etudes Qualitatives*. Editions d'Organisation, Paris.
- Fosso Wamba S., S. Akter., A. Edwards., G. Chopin., D. Gnanzou (2015). "How 'Big Data' Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study." *International Journal of Production Economics*.
- Gandomi, A., and M. Haider (2015). "Beyond the Hype: Big Data Concepts, Methods, and Analytics." *International Journal of Information Management* (35), pp. 137–144.
- Hinds, P. J., and Pfeffer, J (2003). "Why Organizations Don't 'Know What They Know': Cognitive and Motivational Factors Affecting the Transfer of Expertise." in *Sharing Expertise: Beyond Knowledge Management*, M. Ackerman, V. Pipek, and V. Wulf (Eds.), *MIT Press, Cambridge, MA*, pp. 3-26.
- Hung, M., J. Conrad., S.D. Hon., C. Cheng., J.D. Franklin., P. Tang (2013). "Uncovering patterns of technology use in consumer health informatics." *Wiley Interdiscip Rev: Comput Stat*, 5 :6, 432–47.
- Ipe M. (2003). "Knowledge sharing on organizations: A Conceptual framework, *Human Resource Development Review*." 2(4), p. 337-359.
- Isaac, A. (2005). *Conception et utilisation d'ontologies pour l'indexation de documents*. Doctoral thesis, Université Paris IV – Sorbonne.
- Jarvenpaa, S.L., Leidner, D.E. (1999). "Communication and trust in global virtual teams." *Organization Science*, vol. 10:6, 791-815.

- Kankanhalli, A., B.C.Y. Tan., K.K. Wei (2005). "Contributing knowledge to electronic knowledge repositories: an empirical investigation." *MIS Quarterly*, vol 29:1, 113–143.
- Matzler, K., and J. Mueller (2011). "Antecedents of knowledge sharing. Examining the influence of learning and performance orientation." *Journal of Economic Psychology*, vol 32:3, 317-329.
- Nonaka, I. and H. Takeuchi (1995). "The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation." *New York: Oxford University Press*.
- Osterloh, M., and B.S. Frey (2000). "Motivation, Knowledge Transfer, and Organizational Forms." *Organization Science*, vol 11:5, 538-550.
- Prathima, D., K. Parvathy., B. Devi (2011). "Computing Technology Used in Retrieval of Patient's Record." *Proceedings of the National Conference on Innovations in Emerging Technology*.
- Perera, G., A. Holbrook., L. Thabane., G. Foster., D.J. Willison (2011). "Views on health information sharing and privacy from primary care practices using electronic medical records." *International Journal of Medical Information*, vol 80, Issue 2, 94-101.
- Ruggles, R. (1998). "The State of the Notion: Knowledge Management in Practice." *California Management Review*, vol 40:3, 80-89.
- Smith, H.J., T. Dine., H. Xu (2011). "Information privacy research: an interdisciplinary review." *MISQ*, vol 35(4), 989–1016.
- Teigland, R., and M.W. McLure (2003). "Integrating Knowledge through Information Trading: Examining the Relationship between Boundary Spanning Communication and Individual performance." *Decision Sciences*, vol 34:2.
- Thietart, R.-A., al (2014). *Méthodes de recherche en management*, 4ed. Dunod, Paris.
- Wang, S., and A.R. Neo (2010). "Knowledge sharing: A review and directions for future research." *Human Resource Management Review*, vol 20, 115–131.
- Wang, W., L. Shi., A. Yin., Z. Moa., E. Maitland., S. Nicholas., X. Liu (2015). "Primary Care Quality among Different Health Care Structures in Tibet, China." *BioMed Research International*, Vol 2015, Article ID 206709.
- Witherspoon, C.L., J. Bergner., C. Cockrell., D.N. Stone (2013). "Antecedents of Organizational Knowledge Sharing: A Meta-Analysis and Critique". *Journal of Knowledge Management*, Vol. 17 issues 2, 250 – 277.
- Yuwen, S., and X. Yang (2010). "Research on the Clinical Terminology Construction based on SNOMED." *Seventh International Conference on Fuzzy Systems and Knowledge Discovery*.

Appendix

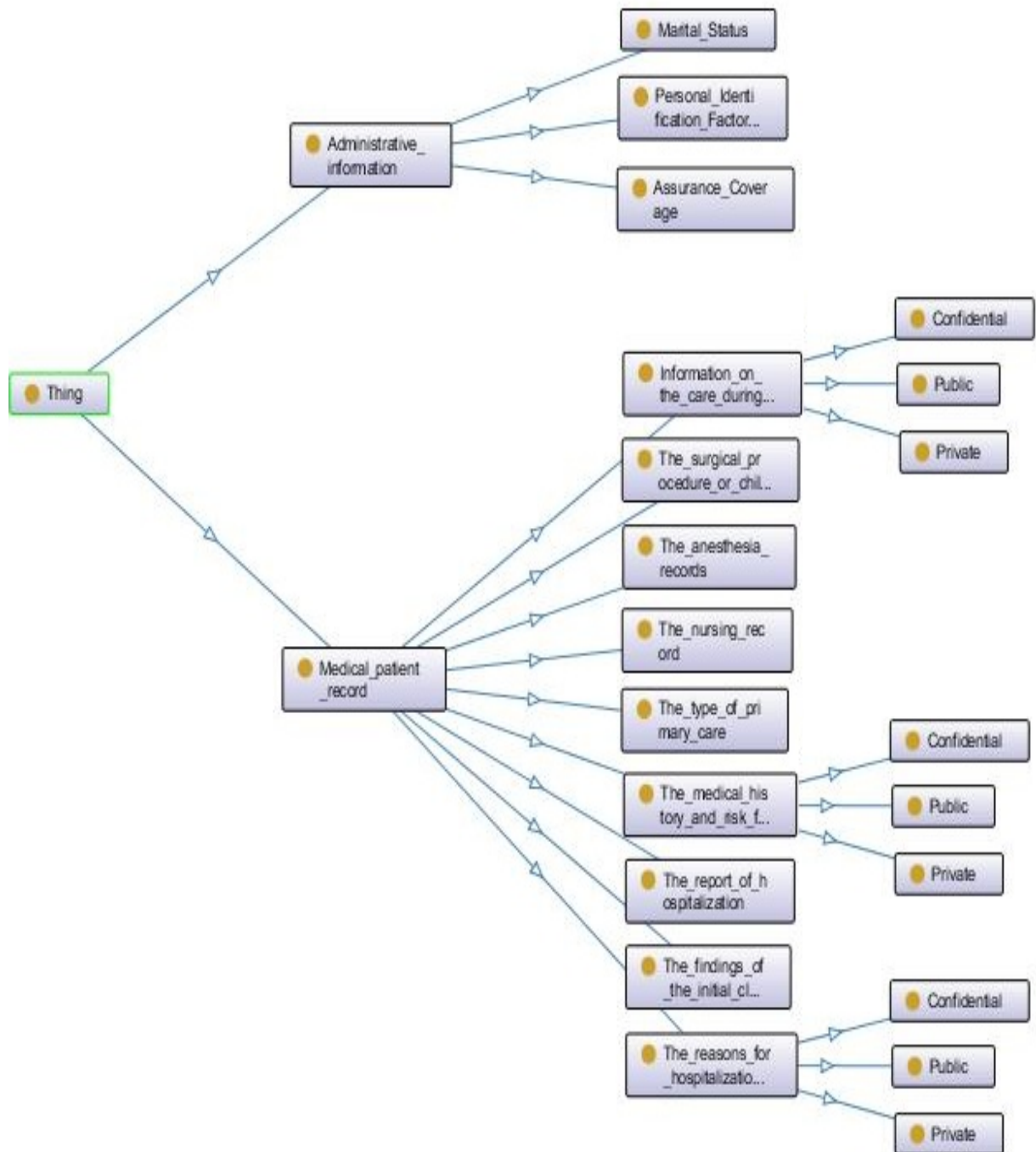


Figure 1. Computerized patient records structure.

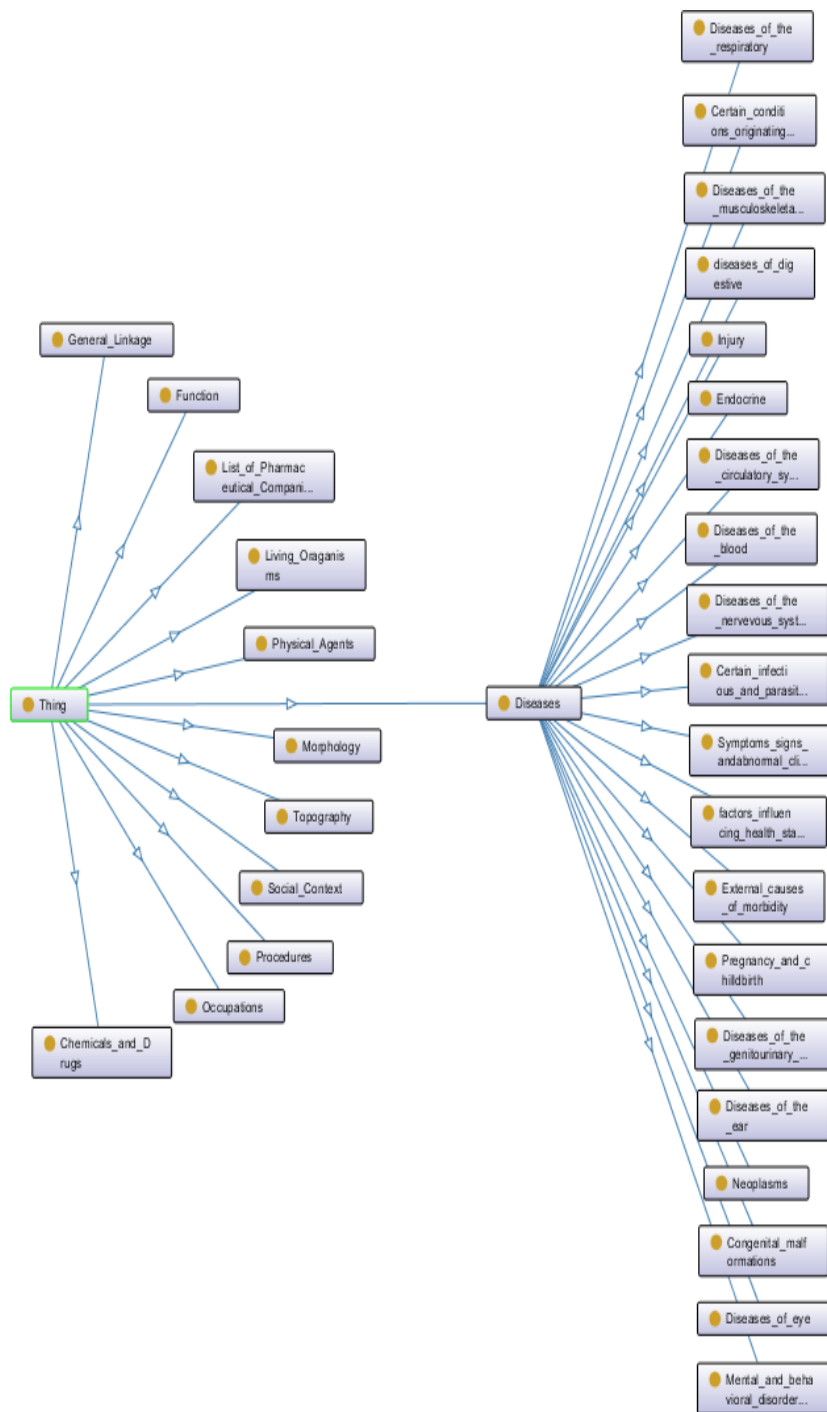


Figure 2. SNOMED taxonomy.

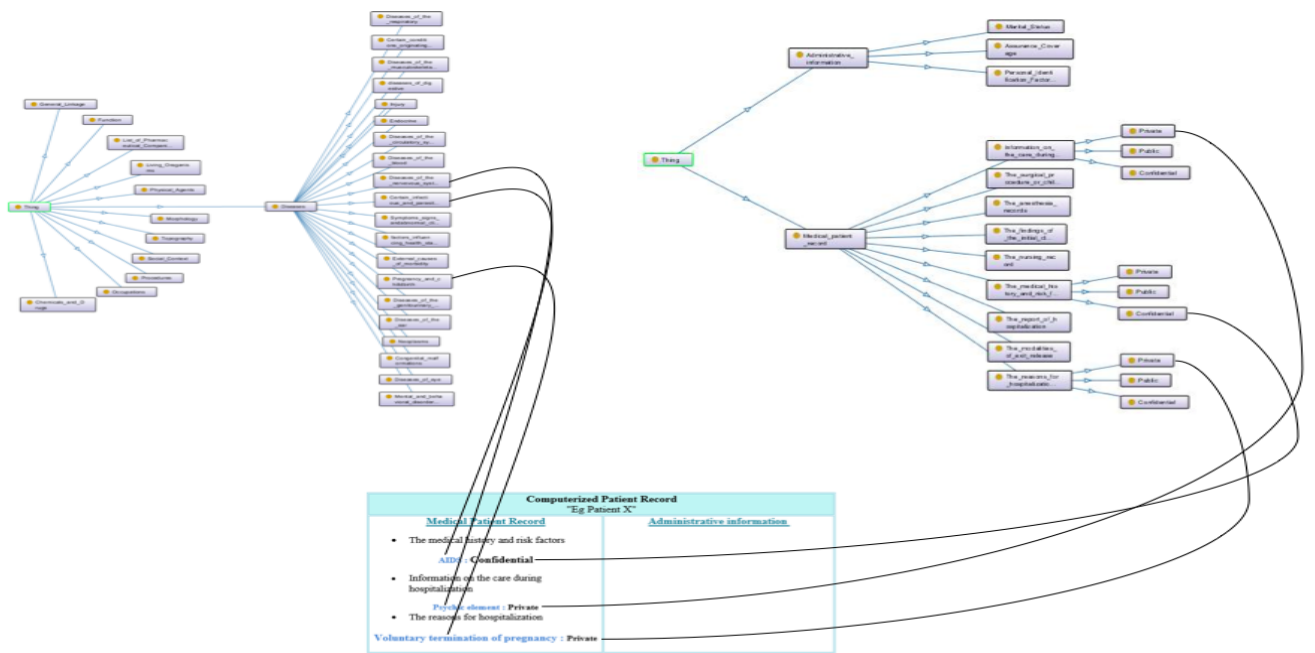


Figure 3. Ontology based electronic patient record annotation.

BIG DATA TECHNOLOGIES: ADDITIONAL FEATURES OR REPLACEMENT FOR TRADITIONAL DATA MANAGEMENT SYSTEMS?

Completed Research

Jemal, Dhouha, University of Tunis, Tunis, Tunisia, dh.jemal@gmail.com

Faiz, Rim, University of Carthage, Tunis, Tunisia, rim.faiz@ihec.rnu.tn

Mahfoudhi, Sami, University of Tunis, Tunis, Tunisia, smahfoudhi@yahoo.com

Abstract

With the data volume that does not stop growing and the multitude of sources that led to diversity of structures, the classic tools of data management became unsuitable for processing and unable to offer effective tools for information retrieval and knowledge management. Thereby, a major challenge has become how to deal with the explosion of data to transform it into new useful and interesting knowledge. Despite the rapid development and change of the databases world, this data management systems diversity presents a difficulty in choosing the best solution to analyze, interpret and manage data according to the user's needs while preserving data availability. Hence, the imposition of the Big Data in our techno-logical landscape offers new solutions for data processing. In this work, we aim to present a brief of the current buzz research field called Big Data. Then, we provide a broad comparison of two data management technologies.

Keywords: Big Data, MapReduce, RDBMS.

1 Introduction

The data is growing at an alarming speed in both volume and structure. With the voluminous data which does not stop growing and the multitude of sources which led to diversity of structures, relational databases which have been proven for over 40 years have reached their limits (Ordonez, 2013), and the classic tools of data management became unsuitable for processing and unable to offer effective tools to find information within massive data and extract value from it. Thereby, as described in (Cuzzocrea et al., 2013), the problem of dealing with the explosion of data has become our major challenge.

As the world becomes more information-driven than ever before, a new technological field has emerged in order to cope with the new requirements for data analysis, information retrieval and knowledge management: the Big Data, as presented in (Sagiroglu and Sinanc, 2013) and (Narasimhan and Bhuvaneshwari, 2014), aims to provide an alternative to traditional solutions database and analysis. Thus, MapReduce (Dean and Ghemawat, 2008) is presented as one of the most efficient Big Data solutions. This framework has found great success in analyzing and processing large amounts of data on large clusters.

Several studies have been conducted to compare MapReduce and Relational DBMS. MapReduce has been presented as a replacement for the Parallel Database Management Systems. However, as proposed in (Stonebraker et al., 2010), MapReduce can be seen as a complement to a RDBMS for analytical applications, because different problems require complex analysis capabilities provided by both technologies.

In this environment of data explosion and diversity, the question arises what technology to use for data analysis and information retrieval, how to benefit the data management systems diversity?

The remainder of this paper is organized as follows. In section 2, we give an overview of the definition and characteristics of big data, presenting some research works that focus on this actual research trend. In section 3, we present the Big Data challenges. Then, we aim to provide a broad comparison of two data management technologies, presenting for each one its strengths and its weaknesses. In section 4, we introduce MapReduce, then in section 5, we describe Relational DBMS. Finally, section 6 concludes this paper and outlines our future work.

2 The Big Data era

Big data presents the next frontier for innovation, competition, and productivity. In this context, we are going to present in this section, the main research work conducted in the Big Data field and to introduce the Big Data characteristics and applications.

2.1 Big Data research work

A strong interest towards the term Big Data is arising in the literature actually. Many research works focus on this actual research trends in the field. In this context, Sagiroglu and Sinanc in (2013), present an overview of big data's content, scope, samples, methods, advantages and challenges, details big data's main component and discusses privacy concern on it. Then, Narasimhan and Bhuvaneshwari in (2014), provides a brief of the buzz-field called Big Data and cover the components of big data from a Hadoop perspective. This study aims to highlight the field's characteristics with two additional dimensions, and to provide a thorough understanding of big data and its various components in the Hadoop framework.

In (Cuzzocrea et al., 2011), open problems and actual research trends are highlighted with the aim of providing an overview of state-of-the-art research issues and achievements in the field of analytics over big data, and extend the discussion to analytics over big multidimensional data. This work presents several novel research directions a rising in this field, which plays a leading role in next-generation Data Warehousing and OLAP research. In (Cuzzocrea et al., 2013), open problems in the field of Data Warehousing and OLAP over Big Data are highlighted. This work aims to present challenges to adopt Data Warehousing and OLAP methodologies with the goal of collecting, extracting, transforming, loading,

warehousing and OLAPing such kinds of data sets, by adding significant add-ons supporting analytic over Big Data.

2.2 Definition, characteristics and applications

The term Big Data, explained by Narasimhan and Bhuvaneshwari, in (2014), was raised the first time by the Gartner office in 2008. It refers to the explosion of data volume and new technological capabilities offered to answer it. Big data can be defined as “a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery, and/or analysis”.

A strong interest towards the term Big Data is arising in the literature actually. Though scalable data management has been a vision for more than three decades and much research has focused on large scale data management in traditional enterprise setting, Big Data brings its own set of novel challenges that must be addressed to ensure the success of data management solutions. This field has received much attention from across the computing and research community, and a lot of work has been done in this context such as (Dean and Ghemawat, 2008; Cohen *et al.*, 2009; Agrawal *et al.*, 2011; Wang and Chan, 2013). The goals of the big data solutions are to meet the new challenges of treating very important volume of structured and unstructured data, located on various terminals.

As presented in (Cuzzocrea *et al.*, 2013), Big Data repositories have two intrinsic factors: (i) size, which becomes really explosive in such data sets; (ii) complexity which can be very high in such data sets. Every day, the amount of data created and manipulated is increasing. All the sectors of activity are affected by this phenomenon. This exponential growth is due to several factors such as trends in the number of users of IT (Information Technology) solutions and data generation by machines. These masses of data bring larger and finer opportunities for analysis as well as new uses of the information. Data today comes from multiple sources such as business transactions and social networks, and in all types of formats: Structured numeric data in traditional databases, and unstructured such as text documents, email, video, audio and financial transactions. Managing, merging and governing both explosion amount and different varieties of data is something many organizations still grapple with.

The web and social networks, whether they are open to all or developed in a professional context, provide kind of opportunities for big data. As described in (Sagiroglu and Sinanc, 2013), McKinsey Global Institute in (Manyika *et al.*, 2011), specified the potential of big data in five main topics:

- Healthcare (clinical decision support systems, analyze disease patterns, improve public health).
- Public sector (discover needs, decision making with automated systems to de-crease risks, innovating new products and services).
- Retail (in store behavior analysis, variety and price optimization, product placement design, web based markets).
- Manufacturing (developed production operations, supply chain planning).
- Personal location data (smart routing, geo targeted advertising or emergency response).

In this context, Big Data approach aims to provide an alternative to traditional solutions database and analysis. Big data solutions add some features to classics DBMSs in order to satisfy new data management needs in a new ecosystem of explosive volume of structured and unstructured data. Actual research trends in the field of Data Warehousing and OLAP over Big Data are rising, such as (Cuzzocrea *et al.*, 2013). The next section presents the general architecture of a data-warehouse.

3 Big Data challenges

Nowadays, Big data and its analysis are at the center of modern science and business. It requires a revolutionary step forward from traditional data analysis. As mentioned in (Sagiroglu and Sinanc, 2013),

the term Big Data is for massive data sets having large, more varied and complex structure with the difficulties of storing, analyzing and visualizing for further processes or results. Big Data, as presented in (Sagiroglu and Sinanc, 2013), is characterized by three main components, the 3 V's: Volume, Velocity and Variety.

1. Volume. It is the first feature brought by the term "big". The size, which can become a real bottleneck from practical applications, refers to the vast amounts of data generated every second. The volume of data stored today is booming. This amount of data that is being collected daily presents immediate challenges for businesses. We can just think of social media messages going viral in seconds, the speed at which credit card transactions are checked for fraudulent activities, or the milliseconds it takes trading systems to analyze social media networks to pick up signals that trigger decisions to buy or sell shares.

2. Velocity. Speed of data in and out; describes the frequency at which data are generated, captured and shared. Growing flows of data must be analyzed in real time to meet the needs of chrono-sensitive processes. Reacting fast enough and analyzing the streaming data is troubling to businesses, with speeds and peak periods of-ten inconsistent. Big Data approach opens the possibility to integrate data streams and generate results or data visualization in (almost) real time.

3. Variety. The volume of Big Data puts data centers in front of challenge: the variety of data. It's not traditional relational data, this data is raw, semi structured or un-structured. In fact, 80% of the world's data is now unstructured¹, and therefore can't easily be put into tables (think of photos, video sequences or social media updates). Big Data is in the form of structured and unstructured data. The structured data types are ready for insertion into a database, while unstructured types have an implicit and irregular structure, and not a fixed pattern (non-relational). With big data technology we can now harness differed types of data (structured and unstructured) including messages, social media conversations, photos, video or voice recordings and bring them together with more traditional, structured data.

In (Narasimhan and Bhuvaneshwari, 2014), we consider two additional dimensions when thinking about big data:

4. Veracity. Accuracy of collected data is a key feature. As mentioned in (Cuzzocrea *et al.*, 2011), very often, data sources, storing data of interest for the target analytic processes, such as web and social networks are strongly heterogeneous and incongruent. Big Data becomes bigger and the multiple sources of big data are ever increasing. So, build confidence in the Big Data represents a significant challenge due to the possibility of inconsistency and abnormality in the Data. Very large data volumes and multiple heterogeneous sources amplify the need for rigor in the collection and crossing data to remove data uncertainty to build confidence and ensure the security and integrity of data.

5. Value. Big Data is gradually transforming organizations around the valuation of information. Big Data approach is designed to achieve the strategic objectives of value creation for the company. With the Big Data approach, the non interesting data when it is taken apart, can take a meaning when considered globally. A big data strategy gives businesses the capability to better analyze data with a goal of accelerating portable growth. Having access to big data, companies generate value from data.

Big data and its analysis are at the center of modern science and business. Inspired by this main motivation, (Cuzzocrea *et al.*, 2011) present a number of open problems and actual research trends related to big data analytics, such as: The data Source Heterogeneity and Incongruence, Filtering-Out Uncorrelated Data, Strongly Unstructured Nature of Data Sources, High Scalability. In this context, a main challenge that has interested the research community and has been the subject of several works such as (Abouzeid *et al.*, 2009; Gruska and Martin, 2010) is: Combining the Benefits of RDBMS and NoSQL Database Systems. It is one of the more relevant features to be achieved by big data analytic systems. As discussed in (Cattell, 2011), it is necessary to combine the benefits of traditional RDBMS database systems and those of the new generation of NoSQL database systems in order to obtain the critical flexibility feature which refers to the property of covering a large collection of analytic scenarios over the same big data partition.

The question is not any more " can Big Data become a relevant competitive ad-vantage? ", but "How can we exploit the opportunities offered by these solutions to optimize our analysis and decision making process? "

4 MapReduce

MapReduce is a programming model developed by Google, which was introduced by Dean and Ghemawat (2008). It was designed for processing large data sets with a parallel, distributed algorithm on a cluster.

MapReduce was created in order to simplify parallel processing and distributed data on a large number of machines with an abstraction that hides the details of the hardware layer to programmers: it hides the details of parallelization, fault-tolerance, locality optimization, and load balancing. Google uses the MapReduce model to deploy large variety of problems such as: generation of data for Google's production web search service, data mining, machine learning, etc.

The MapReduce programming model has been successfully used for many different purposes. These included: parallelizing the effort; distributing the data; handling node failures.

The term MapReduce actually refers to two separate and distinct tasks: Map and Reduce. The mapper is responsible for reading the data stored on disk and process them; it takes a set of data and converts it into another set of data: reads the input block and converts each record into a Key/Value pair. The reducer is responsible for consolidating the results from the map and then write them to disk; it takes the out-put from a map as input and combines those data tuples into a smaller set of tuples.

At first, Google developed their own DFS: the Google File System (GFS). As described in (McClellan et al., 2013), MapReduce tasks run on top of Distributed File Systems (DFS). The distributed storage infrastructure store very large volumes of data on a large number of machines, and manipulate a distributed file system as if it were a single hard drive. The DFS deals with data in blocks. In order to prevent data loss, each block will be replicated across several machines to overcome a possible problem of a single machine failure. So, this model allows the user to focus on solving and implementing his problem.

Nevertheless, the lack is that the MapReduce is independent of the storage system, it can not take into account all the input data for an available index. This explains the critics mainly from the database community. As described in (Gruska and Martin, 2010), the data-base community sees the MapReduce as a step backwards from modern database systems, in view of the MapReduce is a very brute force approach and it lacks the optimizing and indexing capabilities of modern database systems.

MapReduce, the powerful tool characterized by its performance for heavy processing to be performed on a large volume of data that it can be a solution to have the best performance hence makes it very popular with companies that have large data processing centers such as Amazon and Facebook, and implemented in a number of places. However, Hadoop, the Apache Software Foundation open source and Java-based implementation of the MapReduce framework, has attracted the most interest. Firstly, this is due to the open source nature of the project, additionally to the strong support from Yahoo. Hadoop has its own extensible, and portable file system: Hadoop Distributed File System (HDFS) that provides high-throughput access to application data.

Since it is introduced by Google, a strong interest towards the MapReduce model is arising. Many research works aim to apply the ideas from multi-query optimization to optimize the processing of multiple jobs on the MapReduce paradigm by avoiding redundant computation in the MapReduce framework. In this direction, MRShare (Nykiel et al., 2010) has proposed two sharing techniques for a batch of jobs. The key idea behind this work is a grouping technique to merge multiple jobs that can benefit from the sharing opportunities into a single job. However, MRShare incurs a higher sorting cost compared to the naive technique. In (Wang and Chan, 2013), two new job sharing techniques are proposed: The generalized grouping technique (GGT) that relaxes MRShare's requirement for sharing map output. The second technique is a materialization technique (MT) that partially materializes the map output of jobs in the map and reduce phase.

On the other hand, the Pig project at Yahoo (Olston *et al.*, 2008), the SCOPE project at Microsoft (Chaiken *et al.*, 2008), and the open source Hive project 2 introduce SQL-style declarative languages over the standard MapReduce model, aim to integrate declarative query constructs from the database community into MapReduce to allow greater data independence.

5 Relational DBMS

Since it was developed by Edgar Codd in 1970, as presented in (Shuxin and Indrakshi, 2005), the relational database (RDBMS) has been the dominant model for database management. RDBMS is the basis for SQL, and is a type of database management system (DBMS) that is based on the relational model which stores data in the form of related tables, and manages and queries structured data. Since the RDBMSs focus on extending the database system's capabilities and its processing abilities, RDBMSs have become a predominant powerful choice for the storage of information in new databases because they are easier to understand and use. What makes it powerful, is that it is based on relation between data; because the possibility of viewing the database in many different ways since the RDBMS require few assumptions about how data is related or how it will be extracted from the database. So, an important feature of relational systems is that a single database can be spread across several tables which might be related by common database table columns. RDBMS also provide relational operators to manipulate the data stored into the database tables. However, as discussed in (Hammes *et al.*, 2014), the lack of the RDBMS model resides in the complexity and the time spent to design and normalize an efficient database. This is due to the several design steps and rules, which must be properly applied such as Primary Keys, Foreign Keys, Normal Forms, Data Types, etc. Relational Databases have about forty years of production experience, so the main strength to point out is the maturity of RDBMSs. That ensure that most trails have been explored and functionality optimized. For the user side, he must have the competence of a database designer to effectively normalize and organize the database, plus a database administrator to maintain the inevitable technical issues that will arise after deployment.

A lot of work has been done to compare the MapReduce model with parallel relational databases, such as (Pavlo *et al.*, 2009), where experiments are conducted to compare Hadoop MapReduce with two parallel DBMSs in order to evaluate both parallel DBMS and the MapReduce model in terms of performance and development complexity. The study showed that both databases did not outperformed Hadoop for user-defined function. Many applications are difficult to express in SQL, hence the remedy of the user-defined function. Thus, the efficiency of the RDBMSs is in regular database tasks, but the user-defined function presents the main ability lack of this DBMS type.

A proof of improvement of the RDBMS model comes with the introduction of the Object-Oriented Database Relational Model (ORDBMS). It aims to utilize the benefits of object oriented theory in order to satisfy the need for a more programmatic flexibility. The basic goal presented in (Sabàu, 2007), for the Object-relational database is to bridge the gap between relational databases and the object-oriented modeling techniques used in programming languages. The most notable research project in this field is Postgres (Berkeley University, California); Illustra and PostgreSQL are the two products tracing this research.

6 Conclusion

Day after day, zillions of data are generated all over the universe. Many factors contribute to the increase in data volume. The implications of the rise of data generation challenge the needs of data processing, which explains the technological development and diversity of the proposed data management solutions and explosive growth, both in the number of products and services offered and in the adoption of data analysis technologies.

There has been a significant amount of work during the last two decades related to the needs of new supporting technologies for data processing and knowledge management, challenged by the rise of data generation and data structure diversity.

In this paper, we try to detail challenges of the Big Data field. We have presented an overview of big data's characteristics and applications. Although this work has not resolved the entire subject about this trend topic. For this reason, we plan to involve a detailed study on challenges and issues with big data, and to focus on the use cases of Big Data solutions.

References

- Abouzeid, A. Pawlikowski, K. Abadi, D. Silberschatz, A. Rasin, A. (2009). "Hadoopdb: An architectural hybrid of mapreduce and dbms technologies for analytical workloads." In: *Proceedings of the VLDB Endowment*, p. 922-933.
- Agrawal, D. Das, S. and El Abbadi, A. (2011). "Big Data and Cloud Computing: Current State and Future Opportunities." In: *Proceedings of the 14th International Conference on Extending Database Technology*, ACM, p. 530-533.
- Cattell, R. (2011). "Scalable SQL and NoSQL Data Stores." *ACM SIGMOD Record* 39 (4), 12-27.
- Chaiken, R. Jenkins, B. Larson, P.A. Ramsey, B. Shakib, D. Weaver, S. and Zhou, J. (2008). "Scope: Easy and efficient parallel processing of massive data sets." In: *Proceedings of the VLDB Endowment* 1 (2), p. 1265-1276.
- Cohen, J. Dolan, B. Dunlap, M. Hellerstein, J.M. and Welton, C. (2009). "MAD Skills: New Analysis Practices for Big Data." In: *Proceedings of the VLDB Endowment* 2 (2), p. 1481-1492.
- Cuzzocrea, A. Bellatreche, L. and Song, I.Y. (2013). "Data Warehousing and OLAP over Big Data: Current Challenges and Future Research Directions." In: *Proceedings of the sixteenth international workshop on Data warehousing and OLAP*, ACM, p. 67-70.
- Cuzzocrea, A. Song, I.Y. and Davis, K.C. (2011). "Analytics over Large-Scale Multidimensional Data: The Big Data Revolution!" In: *Proceedings of the ACM 14th international workshop on Data Warehousing and OLAP*, ACM, p. 101-104.
- Dean, J. and Ghemawat, S. (2008). "Mapreduce : Simplified data processing on large clusters." *Communications of the ACM* 51 (1), 107-113.
- Gruska, N. and Martin, P. (2010). "Integrating MapReduce and RDBMSs." In: *Proceedings of the 2010 Conference of the Center for Advanced Studies on Collaborative Research*, IBM Corp., p. 212-223.
- Hammes, D. Medero, H. and Mitchell, H. (2014). "Comparison of NoSQL and SQL Databases in the Cloud." *Southern Association for Information Systems (SAIS) Proceedings*, 12.
- Manyika, J. Chui, M. Brown, B. Bughin, J. Dobbs, R. Roxburgh, C. and Byers, A.H. (2011). "Big data: The next frontier for innovation, competition, and productivity." McKinsey Global Institute.
- McClean, A. Conceição, R.C. and O'Halloran, M. (2013). "A Comparison of MapReduce and Parallel Database Management Systems." In: *ICONS 2013, The Eighth International Conference on Systems*, p. 64-68.
- Narasimhan, R. and Bhuvaneshwari, T. (2014). "Big Data A Brief Study." *International Journal of Scientific & Engineering Research* 5 (9).
- Nykiel, T. Potamias, M. Mishra, C. Kollios, G. and Koudas, N. (2010). "Mrshare: sharing across multiple queries in mapreduce." In: *Proceedings of the VLDB Endowment*, vol. 3 (12), p. 494-505.
- Olston, C. Reed, B. Srivastava, U. Kumar, R. and Tomkins, A. (2008). "Pig latin: a not-so-foreign language for data processing." In: *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, ACM, p. 1099-1110.
- Ordonez, C. (2013). "Can we analyze big data inside a DBMS?." In: *Proceedings of the sixteenth international workshop on Data warehousing and OLAP*, ACM, p. 85-92.
- Pavlo, A. Rasin, A. Madden, S. Stonebraker, M. DeWitt, D. Paulson, E. Shrinivas, L. and Abadi, D.J. (2009). "A comparison of approaches to large scale data analysis." In: *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data*, ACM, p. 165-178.
- Sabàu, G. (2007). "Comparison of RDBMS, OODBMS and ORDBMS." *Informatica Economic*.
- Sagiroglu, S. and Sinanc, D. (2013). "Big Data: A Review." In: *Collaboration Technologies and Systems (CTS), 2013 International Conference on. IEEE*, p. 42-47.

- Shuxin, Y. and Indrakshi, R. 2005. "Relational data-base operations modeling with UML." In: *Proceedings of the 19th International Conference on Advanced Information Networking and Applications*, p. 927-932.
- Stonebraker, M. Abadi, D. DeWitt, D.J. Madden, S. Paulson, E. Pavlo, A. and Rasin, A. (2010). "Mapreduce and parallel dbms : friends or foes ?." *Communications of the ACM* 53 (1), 64-71.
- Wang, G. and Chan, C.Y. (2013). "Multi-Query Optimization in MapReduce Framework." In: *Proceedings of the VLDB Endowment, 40th International Conference on Very Large Data Bases* 7 (3).

Open Innovation Strategies and Web Technologies

Mokhtar Amami, Department of Management & Economics, Royal Military College of Canada, Kingston, ON, Canada, amami-m@rmc.ca

Abstract:

The last decade have seen a rapid increase in academic studies and business practices of Open Innovation (OI). A key reason for this has been that OI embodies potentialities of expanding markets through external sourcing of innovation and minimize internal innovation costs. Moreover, it can capture enormous knowledge dispersed around the globe in order to buy/license inventions/Innovations or processes from other businesses. The article aims at characterizing OI strategies. We identify two fundamental dimensions of OI: Innovation uncertainty/complexity and Innovation clock-speed. Combining these two dimensions a conceptual framework of managing OI strategies is proposed. Four distinct typologies or archetypes OI strategies can be obtained: Enterprise network OI strategy; learning & experimentation based OI strategy; collaborative OI strategy; and partnership & alliances OI Strategy. We characterize each one and suggest few insights.

Keywords: Open Innovation, web technologies, strategy, coordination mechanisms, production systems, gatekeepers

Introduction

The Global and dynamic competitive business environment is forcing private and public organizations to look beyond their traditional boundaries for new ideas, new innovations and reliable and innovative network of suppliers. During the last decade Open Innovation (OI) has been the focus of academia and professionals due to their ever-increasing role in time based competition, customer satisfaction of tomorrow needs, shrinking innovation cycles and rising customer expectations.

OI is a relatively a new paradigm shift for business enterprises (Chesbrough, 2003; 2007a, 2007b, 2010, 2011; Enkel et al., 2009) that aims to seize the potentialities of the web and abandon innovation secrecy paradigm to a kind of paradigm of knowledge sharing. The ultimate goal is to expand markets for external use of innovation (Chesbrough 2006), reduce costs of internal innovation and capture the widely distributed knowledge in order to buy/license inventions/Innovations or processes from other businesses.

Although open innovation is attracting more and more researchers, the main problems, however, of creating ideas and sharing them along the supply chain are hard to resolve. In the same token, scanning and capturing new technologies through a web of inventors and startups, or any other channels that can be used as the basis for technology transfer, internal development and joint development exploration and exploitation are even much harder. Granted, a vertical and stable supply chain “à la Toyota, BMW or Walmart” makes the transfer and joint development relatively easy. In fact, the pivot (Toyota, etc.) in this kind of vertical and stable supply chain control the microcosm of process innovation from suppliers to consumers. Suppliers’ involvement in new product development in this type of supply chain is well known and well documented. However, they are still underused and their innovation process is far from integrated in an open innovation process model.

The open innovation process model is much more challenging than the ecosystem of the vertical and stable supply chain. The pitfalls of this rising model are somewhat numerous. The new open innovation process model requires new management competencies for managing fuzzy boundaries, inter-organizational relationships, inter-organizational constraints and bottlenecks. Second, it requires the development inter-organizational information systems to provide visibility. Third, it requires developing trust or at least a kind of understanding to build and nurture collaboration and access the global pool of talent. Finally, it requires new measures to evaluate and share organic value creation

The main research question is how to develop web based OI strategies to scan and channel ideas, knowledge and innovations from varieties of actors distributed around the globe and integrate them in an open innovation process that expand markets through Licensing/Buying, Spin-off and Joint Development?

This research paper is not intended to provide specific answers or solutions to this under-explored question. It is not aiming to build a conceptual model and prove a variety of hypotheses. Its intent is somewhat much simpler and modest. It aims at providing a framework in order to help structure ideas, thinking and discussions related to challenges, opportunities and consequences created by the rise of web technologies and their impacts on open innovation. First, through literature review the paper highlights the fundamental problems of management of innovations in closed innovation ecosystem (closed

innovation) and problems of management of innovations in an environment of sticky information. After discussing briefly OI drivers, we turn to highlighting key central problems of OI. Then, a framework is developed and four archetypes of OI strategies are uncovered. Finally few insights of OI practices are generated and Finally, we conclude with few insights and a brief discussion of possible future research.

2. Background

Management of innovation technology is the study of complex linkages among three central concepts: Innovation, management and technology. Technology refers to the theoretical and practical knowledge, skill, and artifacts that can be used to develop products and services as well as their production and delivery systems. Management refers to the management of idea generation, conversion and diffusion. Innovation refers to “the development and implementation of new ideas by people who over time engage in transactions with others within an institutional order” (Van de Ven, 1986). Management of innovation technology is central to technological progress, economic growth and society development.

Literature on innovation is relatively bulky (Rogers, 1995; Tornatsky & Fleischer, 1990; Tushman & Rosenkopf, 1992; Tushman & Anderson, 1991; Moore, 2002; Christensen, 2003; Christensen & Overdorf, 2000). However, research on “innovation has been narrowly defined on the one hand, and technically oriented on the other” (Van de Ven, 1986). Relatively little of it focuses on the role of interorganizational relationships, environment and particularly on information communication technology (ICT) channels, namely web technologies and applications. Despite this voluminous and extensive research on the determinants of innovation in an organizational setting, the results seem confusing and inconclusive. From Schumpeter (1942) to Thompson (1967) –Rogers (1995) -Van de Ven (1986) - Tornatsky & Fleischer (1990), fundamental theoretical constructs focused mainly on innovation at the level of one organization. However, during the last decade, the dynamic environment is shaping new organizational forms: extended organizational forms. These new organizational arrangement is increasingly attracting researchers (Bensaou & Venkatraman, 1995; von Hippel, 1994) to focus on interorganizational relationships as a medium to enhance innovation through technology transfer, resource exchange and shared learning. Three main reasons seem explain the “sudden” interest of researchers. First, the traditional innovation model (closed innovation - pure hierarchy) appears to have limited power to explain the managerial problems and the implications of technological flows among organizations on one hand and to ensure competitive sustainability. Moreover, the development and deployment of ICT and particularly web-technologies linkages are reducing coordination costs and helping business enterprises and organizations in general to speed up products to market, seize opportunities and increase the value added along the value chains by leveraging each other’s strengths. The use web-technologies can carry rich and intense information among organizations engaged in exchange. Depending on the level of complexities, web-technologies can support the management of innovation value chain (idea generation to conversion to diffusion) across organizational boundaries. To survive in a more and more dynamic environment, organizations are rushing to: 1) search globally for opportunities and resources; 2) focus on core competencies and mutually beneficial longer term outside

relationships; and 3) reap the benefits of global knowledge sourcing by opening the R&D field replacing secrecy by information and knowledge sharing. This new environment will have a major impact on technology diffusion and technology transfer. In summary, because of environment competitiveness, web-technologies development, the global dispersion of innovation competencies, clock speed technological change, process and product complexity and uncertainty of the innovation process, we are seeing more and more organizations (particularly business enterprises) use web-technologies and applications to support their internal innovation capabilities and resources. The strategic intent of this opening R&D is to expand markets (Chesbrough 2006), and capture distributed ideas, knowledge and technologies in order to buy/license inventions/Innovations or processes from other businesses.

3. Open Innovation Drivers

The intensification of global competition combined to the shrinking of innovation cycle are making innovation more and more risky and costly. Facing this competitive environment, organizations and business enterprises in particular are increasingly opening their innovation processes and collaborating across different industries and countries with varieties of innovator actors (e.g.; global academia, researchers in developed countries; researchers in emerging markets; scientists in different industries, retirees, individual networks, dispersed start-ups and laboratories etc.). What drives these global open innovation networks? Many business enterprises find it very difficult to do cost-effective innovation. The global talent pool, cheaper and far “hungrier” is viewed by businesses as an opportunity to accelerate innovation pace and support their business model to make smooth transition toward satisfying tomorrow needs and ensure sustainability. This opening allows organizations to enlarge their peripheral vision which is constrained by the tyranny of actual served customers. The fast cycling and the more and more demanding customer are also pushing businesses to harness tremendous amount of knowledge and expertise dispersed around the globe. Time-to-market will significantly increase if the R&D process is sped up by having the organization connect with external innovator that has developed the technology or is further ahead.

The event of Internet technology and the rise of web-technologies are both the source of new paradigm shift where knowledge is more and more socially constructed and learning and innovation is more and more collective. This new paradigm contrasts radically with the old paradigm characterized by protection and secrecy. It promote collective innovation using collective intelligence (Wang, 2014). Risk reduction through risk sharing is another driver of open innovation. Organizations that combine internal knowledge and innovations with a thorough understanding of outside knowledge innovations help to reduce innovation risk. Finally, open innovation allows businesses to leverage the development and the commercialization of the new product development process. For example, established business enterprises can use open innovation to leverage innovations to take-off full-scale. They can spin out technology or Intellectual Property (IP) that has proved to be outside the core business.

Table 1 contrasts between closed innovation and open innovation and highlights its major drivers.

| Closed Innovation | Open Innovation |
|---|--|
| Enterprises hires the greatest number of talented people for their R&D | Around the globe the number of talented of people working in the same field is 10, 50, 100, 1000, 10000 times greater. |
| Opaque. Protection of IP and Secrecy of idea/ inventions/ innovations | Licensing/buying, spin off and knowledge sharing |
| Individual learning | Collective learning |
| First mover advantage is key to get ahead of the competition | Successful business model is more strategic than to be first to market |
| Control of the whole innovation process excluding competitors profiting from one's idea | Collaboration, knowledge sharing and collective learning create a microcosm of innovations through licensing/buying/spin off innovations and technologies to enhance business model. |
| Innovations are internally business supportive | Innovations are externally business supportive |
| Economies of Scale | Economies of Scope |
| Conflict and competition | Trust and collaboration |

Table 1: Closed Innovation versus Open Innovation

4. Open Innovation Challenges

Management of innovation is a challenging area for any organization private or public. Van de Vent (1986), in a seminal article, provides a thorough analysis of management of innovation drawing from the fields of organizational theory and industrial psychology. The authors asserts that “innovation is the development and implementation of new ideas by people who over time engage in transactions with others within an institutional order.” He exhibited four fundamental problems confronting managers in management of innovation. **The first problem** that challenges managers is the management of attention. Form cognitive psychology Van de Ven shows how people that make up organizations are prone to acquiring routines, group-think, and protecting existing practices instead of paying attention to new ideas. Individuals are restricted by their physiological capabilities that limit ability to process large amounts of stimuli. As organizations become more successful at what they do, the less they are able to recognize new opportunities. Strong leadership is required to overcome the routine to recognize the merits of new ideas. **The second problems** that challenge managers are related to the management of ideas. Institutional change is precipitated by disruptive events which threatens the status quo. Van de Vent suggests that novel ideas need to be championed by individuals so they can become an issue for debate eventually gaining influence and succeeding to make change. **The third problems** that challenge managers are related to the managing part-whole relationships: The innovation process has been traditionally viewed as a sequence of separated parts, which can be seen in firms' functional structures (R&D, marketing, design, etc.). Difficulty in managing part-whole relationships stems from the difficult of overcoming ones interpretive schemas that filter perception: being able to think outside the box. **The fourth problems** of management of innovation is related to institutional leadership: an unsupportive organizational context is likely to undermine the benefits of a well-functioning team; this means that innovative firms need support from their leaders to succeed. Management has thus the responsibility to inculcate appropriate value systems and organizational culture that will foster innovation.

Management of innovation across organizational boundaries is much more challenging than management at the organizational level. First this environment requires two key resources: technologies infrastructures and inter-organizational competencies. Although web-technologies are now widely diffused and less costly, nevertheless because of bandwidth problems and the nature of information, these technologies cannot support all types of innovation across organizational boundaries. In an environment characterized by vertical and stable supply chain, web- technologies can significantly support management of innovation. In an environment of open innovation characterized by dynamic and fleeting opportunities, web-technologies cannot fully support innovations scanning and technology transfer. The latter is particularly challenging because of information and knowledge stickiness.

Von Hippel (1994) identified three major factors that can explain information stickiness and make technology diffusion costly and slow: 1) the nature of the information; 2) the amount of information; and 3) the characteristics of the information seeker and the information providers.

The nature of information can increase or decrease the cost of technology diffusion. There are two types of information: encoded or explicit information and tacit or implicit information. Encoded information is generally related to a process or a product where technical information can easily be transferred. Because design concepts that embody operations principles are mature, well framed and explicitly documented and written, the information is not costly to transfer or not sticky (von Hippel, 1994). However, information related to component knowledge « knowledge about each of the core design concepts and the way in which they are implemented in a particular component - or an architectural knowledge « knowledge about the ways in which the components are integrated and linked together into a coherent whole » can be tacit or implicit or, to use again the von Hippel term, sticky, and thus, costly to transfer. For a variety of processes and products, the knowledge and skills required are not « frozen into the design of machines, but largely stored in the minds of men » (Simon, 1982). Much of the knowledge and expertise used to operate or to solve problems is tacit and requires investments, and individual and organizational learning. Even in the traditional industrial system, the technology developed elsewhere is not free for any organization. Organizations cannot borrow freely the technology. They have to develop their own skills and make their own investment to acquire that sticky part of the technology (Pavitt, 1987). Consequently, information stickiness can be a barrier to technology transfer and diffusion. Individuals and firms alike have to develop their absorptive capacity with respect to new and outside technology in order to reap the potential benefit of the transferred technology. The amount of information can increase the cost of technology diffusion. Technology diffusion may vary according to the amount of information required to implement the technology. Because process and product complexity varies, one may need a very large amount of information to hedge against the contingencies. If we cannot anticipate the exact type and amount of information needed to implement the technology, we require a large amount of information about the use environment of the technology. Consequently, in order to reduce the cost of transferred information, firms have to acquire prior knowledge and make indigenous research. These research activities can help individuals and firms to anticipate the kinds of problems that can arise following the technology transfer. To support our arguments one should review Pavitt (1987),

Cohen & Levinthal (1990) and specially Evenson & Kislev (1975). The characteristics of information seeker and information provider. Technology diffusion cost can vary along the type of the project, firm localisation, product and process type and readiness of the information seeker and the behavior and commitment of the information provider. Teece's (1977) study of international technology transfer projects is very interesting for those firms that are going global. « The costs of transmitting and absorbing all unembodied technological knowledge (i.e., information on methods of organizations and operations, quality control, manufacturing procedures, associated information, but not the knowledge embodied in capital goods, blueprints, or technical specifications)... can vary from 2 percent to 59 percent of total project costs » (see von Hippel, 1994, p. 432). Information cost relative to innovation implementation can also vary along the information seeker. Prior knowledge and the attitude of the information seeker as well as the aim of the technology acquisition are crucial for implementation success. The behavior and commitment of the information provider may also influence considerably the technology diffusion cost. In summary technology scanning, transfer and diffusion is not a free or cheap good.

5. Framework for Open Innovation as a new Mode of Organizing Innovation

5.1 Antecedents of OI strategies

There are key central problems that must be analyzed, evaluated and harnessed in order to seize the opportunities of OI strategies. These are absorptive capacity, complementary assets, intellectual property (IP) and control, technology diffusion, networks and the rise of Peer-Production and Network Externalities and Dominant Design.

Absorptive capacity

Absorptive capacity is critical to organizations and business enterprises in particular to open their internal R&D and start transact ideas, scientific and technological knowledge across their frontiers. Organizations must have the capability to scan, identify technological opportunities, assimilate and integrate them in a profitable business model. Cohen & Levinthal (1990) already underlined this capability of absorptive capacity for firms to generate gains from recognizing, assimilating and adopting external knowledge. Recently other researchers have arrived to similar conclusion. Laursen and Salter (2005) stated that networked companies with high scientific and technical competencies tend to favor OI. Silicon Valley (SV) is an illustration of this phenomenon. The performance of SV cluster is the concentration of high scientific and technological skills that materialize in intense knowledge exchange and fast innovations and adoption (Saxenian, 1999; Kenny & Burg, 1999; Castilla et al., 2000).

Absorptive Capacity can be developed by focusing on the following dimensions:

Awareness: Awareness requires the development of specific processes by which an organization scans for weak signal and uncover knowledge and emerging technologies.

Association: Association requires processes by which an organization uncovers the value of new ideas, prototypes, products or emerging technologies

Assimilation: Assimilation requires the design of processes that fit organizational design by which the organization disseminate and debate these new ideas, prototypes, products or emerging technologies in order to create, capture and deliver value to customers.

Adoption and Implementation: Adoption and implementation requires the development of processes by which an organization transacts ideas, products or technologies for competitive sustainability.

Complementary Assets

Complementary assets are those capabilities (apart from those underpinning the technology) that firms need to exploit the technology (i.e. manufacturing, marketing, reputation, complementary technologies, brand names, supply chain networks, etc.).

The main idea, Teece asserts, that the more complementary assets a company possesses, the more likelihood the company will take advantage of sourcing external scientific and technological knowledge. Teece (1998) distinguished two types of complementary assets: freely available and unimportant and tightly held (kept proprietary) and important. The appropriability regime determines the importance of gains that companies can derive from external knowledge sourcing. Large companies with large assets pool of complementary assets (CA) and tightly held are more oriented toward sourcing external knowledge. The large pool of CA provides potentialities of integrating external knowledge or buying start-ups or buying even large companies. CA, therefore, are important to commercialize innovations and collaborate with other companies.

Intellectual Property: Appropriability and control

Under the traditional innovation paradigm where secrecy and control of intellectual property (IP) are the driving force for a company's ability to capture innovation value. Under the new paradigm of organizing Innovation (Open Innovation), the risk for loss of IP control increases particularly when innovation exchange flows are with competitors (Greenhalgh & Rogers, 2010).. The dilemma of OI is not easy to overcome if the firm avoids developing a kind of understanding and trust relationships based on adaptability to market and technology shifts and alignment of business interests.

IP control and protection refers to the extent to which technology can be protected from imitation. Control and protection depend on the degree to which scientific and technological knowledge is tacit (non-codified). IP control and protection depend on the kind of appropriability regimes: "tight" (e.g., Coca-Cola recipe) vs "weak" (e.g., standard consumer electronic).

Technology diffusion

Technology diffusion is at the heart of intra-organizational and inter-organizational relations (Teece, 1992; Dyer & Ouchi, 1993). Technology diffuses internally and vertically through the interactions of prime contractors and subcontractors (maker - user). As the primes develop their networks of suppliers and affiliated firms, which in turn re-subcontract, diffusion reaches thousands of firms. These networks ensure access to technology and skills from one firm to another (from prime contractor to subcontractor and vice versa). Technology diffuses horizontally through strategic alliances, partnerships, and private and public joint venture research programs. Technology is diffused through a variety of channels: project teams, dispatch of engineers and managers

to partners (prime or subcontractor), hiring of retired technical staff, written documents and so forth. However, the content of these channels is resilient to measurement. The combination of inter-organizational relations and complexities, and information stickiness, information that is costly to acquire, transfer, and use (Von Hippel, 1994) makes it difficult to evaluate the intangible and tangible technological flows among firms.

As outlined earlier, innovation literature is relatively bulky. However, research on “innovation has been narrowly defined on the one hand, and technically oriented on the other” (Van de Ven, 1986). Relatively little of it focuses on the role of up-stream supply chain, environment and particularly on information technology, channels, namely web technologies and applications. Depending on the level of complexities, electronic networks can support the diffusion of technology. For example, the use of Internet technology and web technologies can support organizational resources and capabilities, which in turn, help introduce new administrative practice and support the introduction of new products and processes. Third, to survive in a more and more dynamic environment, the new firm is rushing to: 1) search globally for opportunities and resources; 2) focus on core competencies and mutually beneficial longer term outside relationships; and 3) outsource those activities that can be performed more quickly and at a lower cost by subcontractors or independent business partner. Firms that follow this path find themselves with minimum productive assets, delayed and in the middle of a web, characterized by numerous exchange transactions and complex linkages. In the aerospace, automobile, computer industry and others, firms are building with their partners a competitive supply chain. Dynamic and innovative firms are pressing one step further and already using the web as a major sourcing channel. This new environment will have a major impact on technology diffusion and technology transfer.

In summary, because of environment competitiveness, electronic network development, dispersion of technical competencies around the global, process and product complexity and uncertainty of the innovation process, we are seeing more and more firms use web technologies and applications to support and enhance organizational capabilities and resources, which are the antecedents of internal and external technology diffusion and usage.

Networks and the Rise of Peer-Production

Metcalfe's Law states a network's value grows proportionately to the square of the number of nodes within the network. Although one can dispute the accuracy of this law, intuitively it can provide the most casual observer with insight into the value of networks. All organizations and individuals recognize today the value of networks. In the domain of OI the power of networks to source scientific and technological knowledge (buying and selling) is paramount. The creation and thoroughly designed networks can support the development of a kind of understanding and ultimately a community of trust. Web technologies based networks enhance inter-organizational relationships and interpersonal relationships (Johnson & Duxbury, 2010). These relationships are fundamental to cross-organizational frontiers resources exchange. This environment of web technologies based networks facilitate the strategic roles of OI community drivers. They are external brokers, internal brokers, technology entrepreneurs (P&G), champions (e.g., CEO of P&G) and Capital Venture/Angel entrepreneurs, embodying competencies

that allow them to manage cross-organizational relationships to facilitate bidirectional flows of scientific and technological knowledge. These drivers/leaders are key in diffusing technology, mobilizing volunteers, organizing social networks (Castella et al., 2000) and in managing issues such as technology path divergence (forking problems) and fragmentation of innovation process (balkanization problems) (Fleming & Waguespack, 2007). OI environment requires reputation and trust from OI community leaders to create and sustain an Open Innovation Ecosystem. The key roles of these leaders allow them to resolve issues through social brokerage in connecting OI actors and technological boundary spanning through a process of scanning, identifying, translating, and relaying scientific and technological knowledge across organizational frontiers (Fleming & Waguespack, 2007). Because of their key roles as organizational frontiers spanners, IO community hold a higher degree of trust. Because they have an early access to information and knowledge, they have enormous impact on technology diffusion and control. They can develop and implement varieties of strategies to different groups in an effort to hedge on alternative development. Case studies (e.g., P&G, Peugeot, PG, Coca Cola, etc.) and literature review seem to suggest companies that invest in *technology gatekeepers* not only enhance the flows of scientific and technological knowledge but can immerse themselves in a OI ecosystem that support strategic technological sustainability. Web technologies fostered a new mode of production system that helped the emergence of the rise of the commons and new mode of producing and channeling information and knowledge. The old alternative modes of production, Markets and Hierarchies, are characterized by either high coordination costs or production costs. Web technologies support groups of individuals collaborate on large-scale projects following a diverse cluster of motivational drives and social signals. This web based emerging collaboration is characterized by both low coordination costs and low production particularly when the object of production is information or culture. Because of web technologies and the wide diffusion of computers, communications capabilities and increasingly mobile technologies, OI community will play in the future a major role in scientific and technological knowledge exchange among organizations.

The rise of the commons favor peer-participation where community membership is characterized by *anti-credentials* and there is no *a priori* selection criteria for participation (Bauwens, 2006). Low barriers to participate increase the pool of participants and increase the probability of solutions to problems and particularly the new flow of ideas. Moreover, the lower costs (almost null) associated with peer production allow OI community to play fully their roles without organizational financial constraints. Web technologies based peer production foster diversity of theories, ideas, and perspectives and consequently improve quality outcomes through interactions in innovation open ecosystem.

Peer production system can be characterized by independence, pluralism, representation, decentralized decision-making and autonomous participation (O'Mahony, 2007). This environment of production can provide scientific and technological contribution to firms that invest in gatekeepers.

Network Externalities and Dominant Design

Network externalities effect occurs when open innovation increases the participation of users which increases the value of products for more users. On the demand side, the telephone service illustrates the network externalities. The benefit that people get from the telephone service depends on the extent to which other people also use this service. In other words, network externalities effect happen when “the attractiveness of a product to customers increases with the use of that product by others” (Fisher & Rubinfeld, 2000). Firms with dominant design (standard system) tend to gain enormously from direct network externalities (when an increase in the size of a network increases the number of others). They tend also to gain from indirect network externalities (when an increase in the size of a network expands the range of complementary products available to the users of the network). The more people who adopt the same standard system, the more services and applications the user can access, and so the greater the value of that system to each individual user. On the supply side the firm with the largest network tend to achieve increasing returns to scale because the cost of developing and maintaining the network can be spread over a large and increasing growth of the system or product.

In OI environment network externalities or network effects has a profound impact. The dominant design which is the source of network effects tend to favor an open innovation ecosystem in order to increase the direct and indirect effects. The more users (organizations or individuals) interact with each other, the more ideas and knowledge creation to improve the system and consequently the more the network will increase in value. The increase in the size of network will tend to create an environment of creation of new ideas and knowledge that benefits the creation of more complementary products.

5.2 Framework of Managing OI Strategies

OI embodies buying and selling external knowledge. The fundamental assumption that guides the practice of large firms and start-ups is by opening up the R&D and move away from secrecy, organizations can tap into enormous potentialities of ideas, inventions, innovations and also finding use for “dormant” patents and non-performing products or technologies through selling them or co-exploitation (Chesbrough, 2003, 2006, 2011). What kind of OI strategies should be formulated to reap the potentialities of the new OI ecosystem?

To develop a strategic OI management framework we considered several characteristics that might be useful in classifying OI strategies. At its most fundamental motivation however, organizations are opening-up R&D because of the increasing innovation uncertainty and complexity on one hand and because of the innovation clock speed on the other. Uncertainty is inherent in every innovation process. This uncertainty takes different and concomitant forms such as technological uncertainty, market uncertainty, social and political uncertainty and non-intended consequences uncertainties.

Coping with the rate of change in the world today consumes much attention of organizations and seniors. Markets, technologies, and competitors all move more quickly than a decade ago and twice quickly than two decades ago and may be four times more quickly than three decades ago. Central to this fast speed is innovation. Time-to-market based competition cannot be achieved without clock-speed innovation. The internal

innovation clock-speed is not fast enough to create new processes, new products and services. Because markets are moving faster than internal innovation clock-speed, rivals, new entrants, substitutions and other factors, companies may shift from offensive mode to defensive posture. To avoid this situation and technological and market disruptions, business enterprises are opening their R&D to source external scientific and technological knowledge.

Summing-up one can identify two dimensions that can be used to characterize OI strategies: Innovation uncertainty/complexity and Innovation clock-speed. Combining these two dimensions a conceptual framework of managing OI strategies is proposed. Four distinct typologies or archetypes OI strategies can be obtained: Enterprise network OI strategy; learning & experimentation based OI strategy; collaborative OI strategy; and partnership & alliances OI Strategy. The relationship of these four configurations is illustrated in Figure 1.

Before describing and discussing each typology, it is imperative to stress the following elements for better characterization of these typologies. The elements that best characterize typologies are: coordination mechanisms, strategic intent, gatekeepers, production system and business case that illustrate the typology. Among these elements coordination mechanisms (communication & control) is central to OI. The literature stressed three main coordination mechanisms (Mintzberg, 1979, Galbraith 1973): 1) mutual adjustment, 2) standardization, and 3) shared variable. However, web technologies (web 2.0 and their applications and web 3.0) (Chui et al., 2009; Flat World Business, 2014) can serve as an enabler and supporter of inter-organizational relationships and particularly in information and scientific and knowledge sharing. By allowing rapid information flows at much lower costs, web-technologies enable interactive relationships that otherwise might have been desirable but impractical. In OI environment web-technologies play a central role as a primary coordination mechanisms.

Enterprise Network OI Strategy (L-L Environment)

In an environment characterized by relatively low innovation complexity and uncertainty coupled to low innovation clock-speeds (Quadrant L-L – Figure 2), organizations tend to adopt an OI strategy that tends to network primarily their own **stable** supply chain where they play the role of pivot. The case of Coca-Cola products offers an excellent illustration of this type of OI strategy. Coca Cola processes and products have in general low level of technological complexity and characterized by stable technology cycle. Because of appropriability regime that characterize Coca Cola products, IP does not create cross-organizational relationships problems. However, this type of company is not sealed completely from technological progress. Today branding is not limited to the unique characteristics of the product, but go beyond to incorporate issues such as environment and corporate social responsibility (e.g., suppliers' relationships, contribution to local community). In this kind of context, green supply chain (waste management, water usage, packaging etc.) becomes fundamental to strategic sustainability. Granted, companies operating in the quadrant L-L, the impact of innovation on processes and products are relatively low. However, OI within the frontiers of the business enterprise and across its vertical supply chain is important. For that reason, they must use the application of Web 2.0 (Wagner & Jiang, 2012; Ooi et al.,

2011) which is Enterprise 2.0 (see Annex 2) to capitalize on internal ideas and knowledge and across their stable supply chain (vertical) to improve existing activities. Further, this type of business enterprise can use Enterprise 2.1 (see Annex 2) to interact with consumer and outside communities to gather ideas and knowledge particularly in the domain related to green supply chain. For strategic sustainability, companies in the Quadrant L-L can support OI by adopting Enterprise 2.0 and eventually Enterprise 2.1 based coordination mechanisms.

Coca-Cola

Coca Cola famous for keeping the recipe of its drink secret may not need OI to come up with new products or design new technological processes. However, their green supply chain and their Corporate Social Responsibility (e.g.; unemployment) is leading them to seek new ideas and knowledge to create better living for society. Combined to provide solutions to real society problems, the restricted opening of Coca-Cola aim at improving their corporate image on one hand and support their strategic sustainability on the other. The Web 2.0 and Enterprise 2.1 are used to create a network that support idea generation, interaction and debating to help converge toward solutions.

Collaborative OI Strategy (H-L Environment)

The collaborative Open Innovation strategy (H-L Environment- Figure 3) aims at mining opportunities through sourcing external information, scientific and technological knowledge. In order to accelerate internal innovation processes and deliver greater overall value to the market, P&G's "Connect + Develop" model was developed.

P&G CEO and senior managers believe that innovation constitutes the foundation of P&G Model. The company used the new organizing innovation model to create value with extended network of innovators, start-ups, established businesses, customers, and individuals through InnoCentive reaching untapped talent pools (Huston & Sakkab, 2006; Ruiz, 2009). The ultimate goal is satisfy customers and maintain strategic sustainability. P&G new organizing innovation model is customer focused rather than technology focused. Innovation is not limited to the upstream value chain. The innovation is involved along the product value chain in areas such as packaging, shopping experience, after the sale service and the in-home product usage experience. New structure called "Connect and Develop" was created. At its heart two driving forces: 1) Gatekeepers (Research Fellows and technology entrepreneurs) and 2) Enterprise 2.0 and enterprise 2.1.

Enterprise 2.0 and Enterprise 2.1 are both applications of web 2.0. These 2 applications help increase the involvement of employees in establishing and managing data. Moreover, P&G is also allowing its employees to personalize its web portal by adding RSS feeds of news and business information. These web- technologies based service allow the development of relationships with customers, suppliers, online hubs for all the interactions with supermarkets, scientific and technological knowledge providers around the world. Web based collaboration at P&G succeeded to source external knowledge through the creation of a network of weak ties (Ruef, 2002; Castilla et al., 2000) and also because of absorptive capacity, internal talent, and particularly gatekeepers that play a critical role as technological entrepreneurs. Combining the power of web technologies (Gruber, 2008) network and scouts (Technology entrepreneurs), P&G succeeded to hunt for new ideas, inventions and innovations cross organizational boundaries. Another type of gatekeeper that sustain absorptive capacity is the In-House R&D team (Research

Fellows at P&G) responsible for conceptual development and continuous refinement. This type of gatekeeper is a hedge against the risk of losing expertise to innovate.

Knowledge management tools are embedded deeply in the structure, supported by web based technologies such as web platforms like “Innovation Net” and Enterprise 2.0 and Enterprise 2.1 that leverage competencies of gatekeepers (research fellows and technology entrepreneurs), who play central role in the collaborative OI strategy of P&G.

Philips a pioneer and innovator of health and well-being products, adopted a somewhat different collaborative approach to P&G one can call it collaborative OI hybrid strategy (Philips, 2011). OI Philips Research strategy aims at maximizing the benefits of both outside-in and inside-out innovation approaches. This strategy is motivated by innovation stickiness (tacit knowledge), high intellectual proprietary regime and diffusion characteristics. For Philips collaborative innovation is a means to benefit from complementarities assets and complementarities among firms.

Learning & Experimentation OI Strategy (H-H Environment)

In the quadrant H-H (Figure 4) companies leverage the speed and the power of web technologies to manage the flow of ideas, scientific and technological knowledge and integrate them in fast flexible and responsive processes for accelerating understanding and learning of technology uncertainty and complexity in an environment of high-clock speed. In this quadrant companies use pure peer production system and web technologies as coordination mechanisms. They tend to combine knowledge generation through rapid cycles of experimentation and the rapid integration of the information generated. The use of web technologies based pure peer production system and the design of extreme flexible-structured-controlled processes allow companies to generate, document, develop and control activities. The following two cases illustrate the environment of this quadrant.

PSA Peugeot Citroën

During 2011, **PSA Peugeot Citroën** (French car manufacturer) has initiated a collaborative project to design the car of the future. The company developed a collaborative effort with dispersed scientific laboratories around the globe. The launched project yielded a network of OpenLabs that allows collaboration between internal research centers and external research groups. The aim is to take advantage from dispersed ideas and knowledge when connected can create a platform for interactions and produce the car of future that fits for example the environment requirements and customer needs.

Local Motors

Local Motors is a start-up created in 2007 by a former marine, Jay Rogers. Designing a new car requires multiple teams and multiple competencies, experience and huge resources. The use of crowd sourcing can save the start-up time and money. Moreover, by using crowd sourcing, the start-up can gain from dispersed skills and experiences around the global and creating a global product that may speed-up its diffusion and adoption.

The H-H environment is characterized by a relatively low risk of external sourcing (low competitive imitation) and the need for high external knowledge and sophisticated competencies. The use of web-technologies based sourcing external knowledge allows companies to sense the market and somewhat predict the major source of unpredictable change of the market needs.

In the quadrant H-H, it is very difficult to predict tomorrow needs. By combining web technologies and its collective intelligence (Lesser et al., 2012) peer production and the

design of extreme flexible processes, companies can increase knowledge sharing and enhance organization learning in order to react, adapt and design very rapidly new products and new processes to fit market needs. Companies in this quadrant require gatekeepers with high competencies to integrate external scientific and technological knowledge and market needs. Their task consist of reacting continuously to the flow of information and knowledge in a synchronized way.

Partnerships & Alliances OI Strategy (L-H Environment)

The quadrant L-H (Figure 5) is characterized by relatively low innovation clock-speed and relatively high innovation uncertainty and complexity. Companies operating in this quadrant combine hierarchy, market and peer production system. They use Enterprise 2.0, Intranet and Extranet technologies as part of their coordination mechanisms. The gate keepers are externally oriented managers and virtual teams with high competencies in managing inter-organizational relationships.

Companies in this quadrant are Aerospace, pharmaceutical, consumer electronics and automobile manufacturers. They tend to keep internally activities such as design, branding, marketing, but outsource most other activities (assembly included for consumer electronics companies).

OI strategy is based on organizational alliances and partnerships (P&A) focused on creating and *capturing* value along the supply chain. Companies involved in P&A to source external knowledge in order to reduce innovation uncertainty and complexity need particular governance mechanisms and special contracting.

There are three types of governance mechanisms, markets, hierarchies and clans (Ouchi, 1980). For companies operating in L-H environment and engaged in OI strategy based on P&A, the clan mechanism determines the transaction costs (Williamson, 1975, 1979) through a system where contributions and retributions are balanced on the long run. Because of the nature of innovation (characterized by its uncertainty and complexity), clan mechanisms is the proper mechanism. In this environment characterized by performance ambiguity and difficulty of goal congruence among players engaged in sourcing external knowledge, relational contracting (Williamson, 1979) is best.

The environment L-H requires adaptability and alignment. Uncertainty and complexity of the technology require time and adaptation to changing needs caused by shifts in markets. For instance, structural shifts can occur due to economic progress, political and social change, demographic trends, and new emerging technology. Therefore, companies engaged in P&A for external knowledge sourcing must adapt their inter-organizational relationships and their processes to ensure the best outcome.

Alignment is a form of trustworthy relationships where risks, costs and rewards are equitably shared. Misaligned interests and OI strategies among players engaged in P&A is a sure path for P&A failure. If any participating member's interest is not aligned, it can cause conflict of interest and failure of OI strategy implementation and execution. Hence, it is critical that all participating member's interests are aligned so that OI strategy achieves its strategic intent. Finally, alignment will ensure a kind of understanding with minimum of ambiguity and of equivocality related to the purposes and priorities of the P&A.

6. Insights for practicing OI

1. The use of external sourcing, external and internal experimentation integration, and absorbing new ideas and experiences particularly from external sources constitute a real challenge. Not all organizations are ready to manage this variety of information, knowledge and technology flow and integrate it in a competitive business model.
2. We need to develop theoretical foundations for external innovation sourcing. Theoretical constructs such as RBV, TCT, network externalities, and absorptive capacity must be integrated in a coherent explanatory and predictive model.
3. IP issues is problematic (Pisano & Teece, 2007). How protect IP? Have safeguards in place to make sure IP is not accidentally disclosed, or external IP is not used in an illegal way.
4. Web technologies are blessing. However, connection and reach are not so easy to establish. It needs specific ICT infrastructures an applications integrated in organizational processes in order source external knowledge particularly for sticky knowledge.

7. Conclusion:

In review of the current literature it can be seen that external knowledge and expertise contributing to new innovations and new products development is increasing rapidly enabled by the evolution of the web and its collaborative capabilities. Web 3.0 and Web 2.0 and its application Enterprise 2.0 and Enterprise 2.1 has moved innovation from a traditional paradigm characterized by internal knowledge sourcing and secrecy to a new paradigm of organizing innovation (open innovation).

The fundamental assumption underlying open innovation paradigm is to bring external research institutions, customers, suppliers, competitors, retirees and individuals to innovate together through information and knowledge sharing (Ruiz, 2009). However, external sourcing of information and scientific & technological knowledge is challenging. Probably, information stickiness (Von Hippel, 1994) is the most important challenge to external knowledge sharing. Moreover, in order to seize the opportunities of OI strategies, key central problems must be analyzed, evaluated and harnessed. These are absorptive capacity, complementary assets, intellectual property (IP) and control, technology diffusion, networks and the rise of Peer-Production and Network Externalities and Dominant Design.

Through a brief and limited review of open innovation practices by most known and publicized business enterprises, fundamental dimensions for developing a framework for OI strategies are uncovered. Then primary channels and coordination mechanisms used to source external innovations are identified.

The framework highlighted in figure one shows four types or archetypes of OI strategies: Enterprise network OI strategy; Learning & Experimentation based OI strategy; Collaborative OI strategy; and Partnership & Alliances OI Strategy. How in each configuration organizations scan, identify and source ideas and innovations? Gatekeepers with high technological and inter-organizational competencies. What coordination mechanisms organizations are using to collaborate with outsiders like universities, suppliers, customers, global academia, researchers in developed countries; researchers in emerging markets; scientists in different industries, retirees, individual networks,

dispersed start-ups and laboratories? Web technologies are used as one of the primary factor for coordination mechanisms. What production systems are using to generate ideas and technological knowledge? Markets, hierarchies and peer production systems are combined and used. In quadrant H-H, for example, companies tend to use pure peer production system.

The aim of this paper is to build on prior contributions in the domain through the development of OI framework. Four types or archetypes of OI strategies are identified. Moreover, differences among these four OI strategies in terms of coordination mechanisms, gatekeepers, strategic intent and production system are uncovered.

Granted, much remains to be done from theoretical and empirical view. Practices of well-known and publicized such as P&G (2006), Starbuck, Coca Cola, Philips (20011), Peugeot, IBM, Lego, Local Motors, GE, Fiat, Cisco and others is not enough to validate the framework. The next important step is to formulate hypotheses and confront the framework to a more rigorous empirical testing.

8. References

- Bauwens, M.** (2006). *The Political Economy of Peer Production*. Thailand: Payap University and Chiang Mai University.
- Bensaou, M.** and Venkatraman, N. (1995), Configurations of Interorganizational Relationships: A Comparison between U.S. and Japanese Automakers, *Management Sc.*, Vol.41, pp.1471-1492.
- Castilla, E.J., Hwang, H., Granovetter, E. & Granovetter, M. (2000). Social networks in Silicon Valley. Chapter 11 in *The Silicon Valley Edge: A Habitat for Innovation and Entrepreneurship*, edited by Chong-Moon Lee, William F. Miller, Henry Rowen, and Marguerite Hancock. Stanford: Stanford University Press.
- Chesbrough, H. W.** (2011). The Era of Open Innovation. *MIT Sloan Management Review*(Winter), 35-41.
- Chesbrough, H.W.** (2003) Open Innovation: The New Imperative for Creating and Profiting from Technology. Cambridge, MA: Harvard Business School Publishing.
- Chesbrough, H.W.** (2006) Open Business Models: How to Thrive in the New Innovation Landscape. Cambridge, MA: Harvard Business School Publishing.
- Chesbrough, H.W. (2007a) Why companies should have open business models. *MIT Sloan Management Review*, 48, 2, 22–28.
- Chesbrough, H.W. (2007b) Business model innovation: It's not just about technology anymore. *Strategy and Leadership*, 35, 6, 12-17.
- Chesbrough, H.W. (2010) Business model innovation: Opportunities and barriers. *Long Range Planning*, 43, 354-363.
- Chesbrough, H.W.** (2011) *Open Services Innovation: Rethinking Your Business to Grow and Compete in a New Era*. San Francisco, CA: Jossey-Bass Publishing.
- Christensen, Clayton (2003). *The Innovator's Solution: Creating and Sustaining Successful Growth*. Harvard Business School Press.
- Christensen, Clayton M., & Overdorf, Michael.** (2000). *Meeting the Challenge of Disruptive Change*. *Harvard Business Review*, 78(2), 66–76.
- Chui, M., Miller, A., & Roberts, R. P.** (2009). Six ways to make Web 2.0 work. *The McKinsey Quarterly*. Retrieved from [http://www.eclf.org/newsite/source/10-29-09/Six%20ways%20to%20make%20Web%202.0%20work%20\(MKQ%202-2009\).pdf](http://www.eclf.org/newsite/source/10-29-09/Six%20ways%20to%20make%20Web%202.0%20work%20(MKQ%202-2009).pdf)
- Cohen, W.** and Levinthal, D. (1990), "Absorptive capacity: A new perspective on learning and innovation", *Administrative Science Quarterly*, pp. 128-152.
- Di Minin, A., Frattini, F., & Piccaluga, A.** (2010). Fiat: Open Innovation in a Downturn (1993-2003). *California Management Review*, 52(3), 132-159.
- Dyer, J.H.** and Ouchi, W.G. (1993), "Japanese - Style Partnerships: Giving Companies a Competitive Edge", *Sloan Management Review*, Fall Issue, pp. 51-63.

- Enkel, E., Gassmann, O., & Chesbrough, H. (2009). Open R&D and open innovation: exploring the phenomenon. *R&D Management*, 39 (4), 311-316.
- Flat World Business. (2014). *Web 1.0 vs Web 2.0 vs Web 3.0 vs Web 4.0 – A bird's eye on the evolution and definition* | Flat World Business. Retrieved June 25, 2014, from <http://flatworldbusiness.wordpress.com/flat-education/previously/web-1-0-vs-web-2-0-vs-web-3-0-a-bird-eye-on-the-definition/>
- Fleming, L., & Waguespack, D. M. (2007). Brokerage, Boundary Spanning, and Leadership in Open Innovation Communities. *Organization Science*, 18(2), 165-180.
- Galbraith, J. (1973), *Designing Complex Organizations*, Addison-Wesley.
- Gassmann, O. (2006). Opening up the innovation process: towards an agenda. *R&D Management*, 36 (3), 223-228.
- Gassmann, O., & Enkel, E. (2004). Towards a Theory of Open Innovation: Three Core Process Archetypes. *Innovation*, 1-18.
- Greenhalgh, C., & Rogers, M. (2010). The Nature and Importance of Innovation. In *Innovation, intellectual property and economic growth*. Princeton: Princeton University Press.
- Gruber, T. (2008). Collective knowledge systems: Where the Social Web meets the Semantic Web. *Journal of Web Semantics*, 6(1), 4-13. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1570826807000583>
- Henkel, J. (2006). Selective revealing in open innovation processes: The case of embedded Linux. *Research Policy*, 35, 953-969.
- Huston, L., & Sakkab, N. (2006). Connect and develop: Inside Procter & Gamble's new model for Innovation. *Harvard Business Review*, 84(3), 58-66.
- Johnson, K. L., & Duxbury, L. (2010). The view from the field: A case study of the expatriate boundary-spanning role. *Journal of World Business*, 45, 29-40.
- Kenny, M. & Burg, U.V. (1999). Technology, entrepreneurship and path dependence: industrial clustering in Silicon Valley and route 128, *Industrial and Corporate Change*, Vol. 8, No. 1
- Lesser, E., Ransom, D., Shah, R., & Pulver, B. (2012). *Collective Intelligence Capitalizing on the Crowd*. Retrieved from IBM Institute for Business Value website: http://www.bicinnovation.com/static/bic/knowledge_base/documents/IBM3.pdf
- Mintzberg, H. (1979), *The Structuring of Organizations*, Prentice Hall, Inc.
- Moore, Geoffrey (2002). **Crossing The Chasm: Marketing and Selling Disruptive Products to Mainstream Customers**. Harper Collins Publishers.
- O'Mahony, S. (2007). The governance of open source initiatives: what does it mean to be community managed? *Journal of Management and Governance*, 11, 139-150.
- Ooi, K., Chong, A. Y., & Tan, B. (2011). Application of Web 2.0 in Supply Chain Management: A Brief Overview. *rends in Applied Sciences Research*, 6. Retrieved from <http://scialert.net/abstract/?doi=tasr.2011.394.399>
- Ouchi, William G. 1980. "Markets, Bureaucracies, and Clans." *Administrative Science Quarterly* 25:129-41.
- Pavitt, K. (1987), "The Objectives of Technology Policy", *Science and Public Policy*, Vol. 14(4), pp. 182-188.
- Philips. (2011). *Philips Research: Open Innovation*. Retrieved Nov 26, 2011, from Philips: <http://www.research.philips.com/open-innovation/index.html>
- Pisano, G. P., & Teece, D. J. (2007). How to Capture Value from Innovation: Shaping Intellectual Property and Industry Architecture. *California Management Review*, 50(1), 278-296.
- Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). NY: Free Press.
- Ruef, M. (2002). Strong ties, weak ties and islands: structural and cultural predictors of organizational innovation. *Industrial and Corporate Change*, 11(3), 427-449.
- Ruiz, P. P. (2009). *Procter & Gamble's Connect and Develop* | *Open Innovate*. Retrieved from <http://www.openinnovate.co.uk/pgs-connect-and-develop/>
- Saxenian, A. (1999). Comment on Kenney and von Burg, "Technology, entrepreneurship and path dependence: industrial clustering in Silicon Valley and Route 128", *Ind Corp Change*, Vol. 8, p 105-110
- Schumpeter, J. (1942), *Capitalism, Socialism, and Democracy*, Harper and Row, New York.
- Simon, H.A. (1982), *Models of Bounded Rationality: Behavioral Economics and Business Organization*, Vol. 2, MIT Press, Cambridge, MA

- Teece, D.** (1998), "Capturing Value from Knowledge Assets: The New Economy, Markets for Know-how, and Intangible Assets", *California Management Review*, Vol. 40, No. 3, pp. 55-79.
- Teece, D.J.** (1977), "Technology Transfer by Multinational Firms: the Resource Cost of Transferring Technological Know-How", *Economic Journal*, Vol. 87, pp. 242-261.
- Teece, D.J.** (1988), "Technological Change and the Nature of the Firm", in G. Dosi, et al. (Eds.), Technical Change and Economic Theory, Frances Pinter, London.
- Teece, D.J.** (1992), "Competition, Cooperation, and Innovation - Organizational Arrangements for Regimes of Rapid Technological Progress", *Journal of Economic Behavior and Organization*, Vol.18, pp.1-25.
- Thompson, J. D.** (1967), Organizations in Action, McGraw-Hill, New York.
- Tornatzky, L.G.** and **Fleischer, M.** (1990), The Process of Technological Innovation, Lexington Books, Lexington, MA.
- Tushman, M** and **Anderson, P.** Managing Through Cycles of Technological Change, in *Research-Technology Management*, May-June 1991
- Tushman, M** and **Rosenkopf, L.** Organizational Determinants of Technological Change, in *Research in Organizational Behavior* 14, 1992, JAI Press
- Van de Ven, A. H.** (1986). « Central problems in the management of innovation". *Management science*, 32(5), 590-607.
- Von Hippel, E.** (1994), "Sticky Information and the Locus of Problem Solving: Implications for Innovation", *Management Science*, Vol. 40(4), pp. 429-439.
- Wagner, C., & Jiang, L.** (2012). Harnessing the power of social media for creativity support: A three-pronged approach. *Knowledge Management & E-Learning: An International Journal*, 4(2). Retrieved from <http://www.kmel-journal.org/ojs/index.php/online-publication/article/viewFile/183/140>
- Wang, K.** (2014). Collective Innovation: The Known and the Unknown. *Collective Intelligence*. Retrieved from <http://humancomputation.com/ci2014/papers/Active%20Papers%5CPaper%2061.pdf>
web.stanford.edu/group/esrg/siliconvalley/docs/siliconvalleyedge.pdf
- Williamson, Oliver.** 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York, NY: Free Press.
- Williamson, Oliver.** 1979. "Transaction Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics* 22: 233-61.

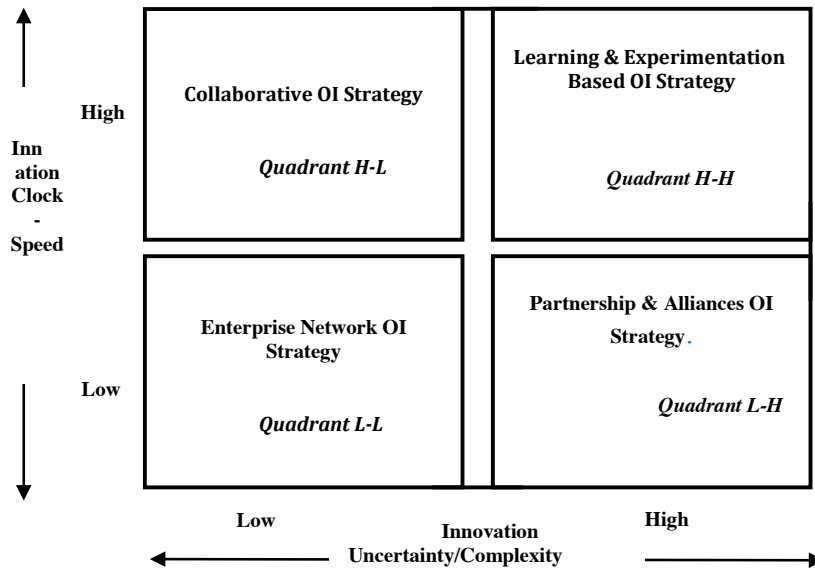


Figure 1: Typology of OI Strategies

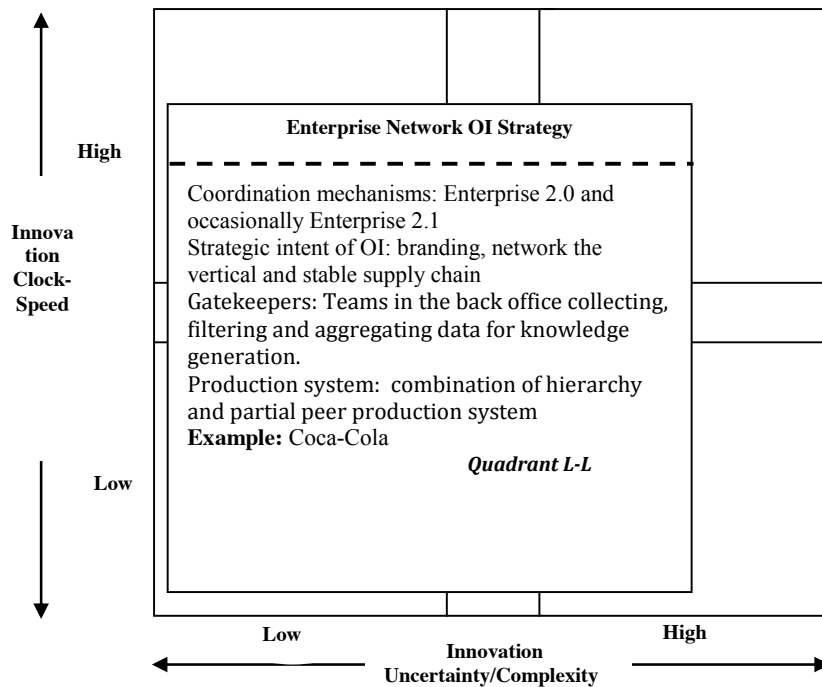


Figure 2: Stable OI Strategy

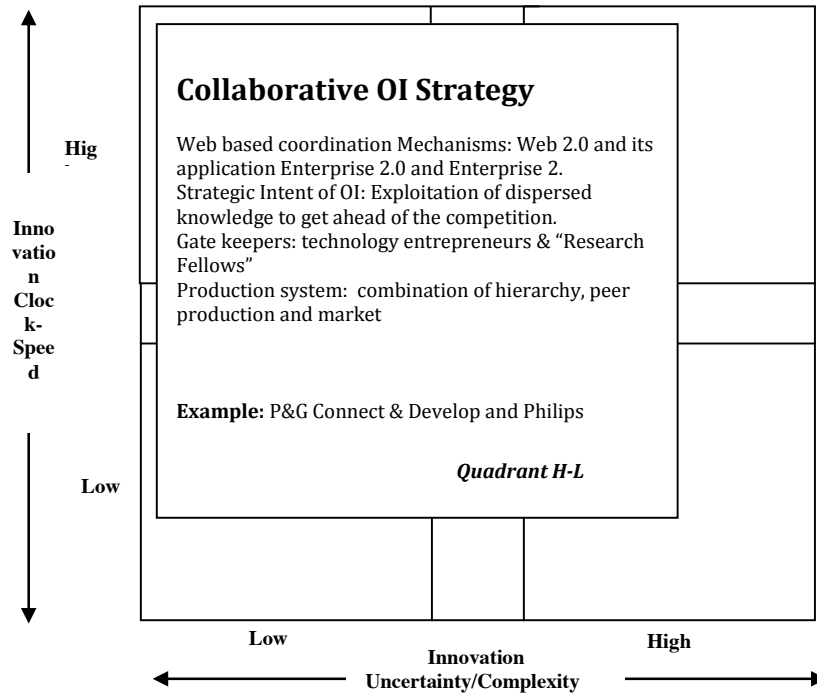


Figure 3: Collaborative OI Strategy

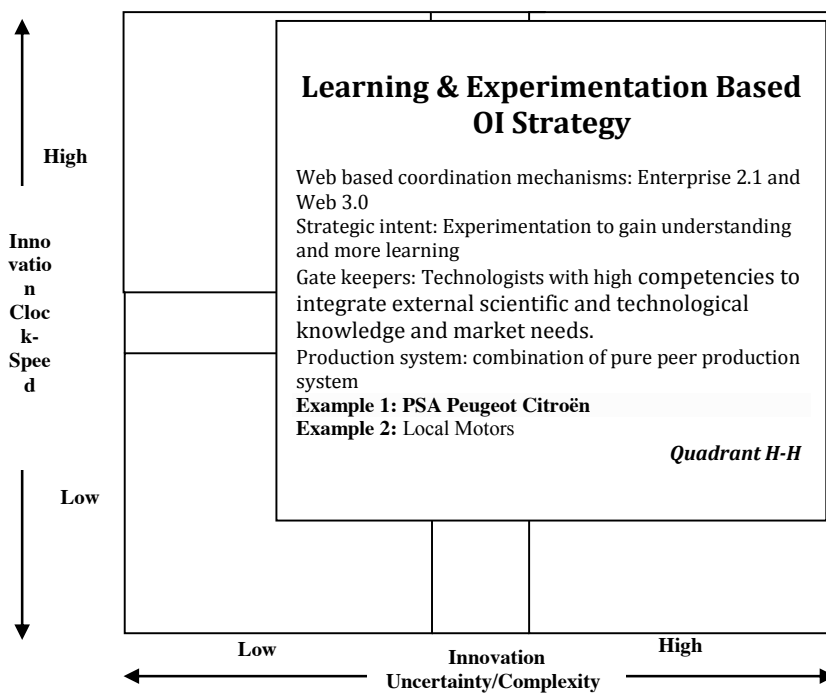


Figure 4: Learning and Experimentation Based OI Strategy

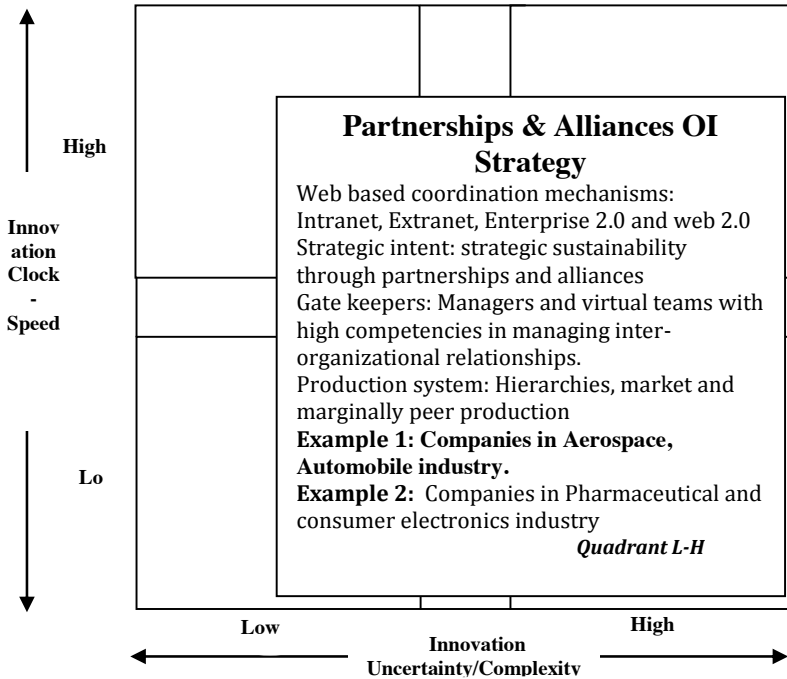
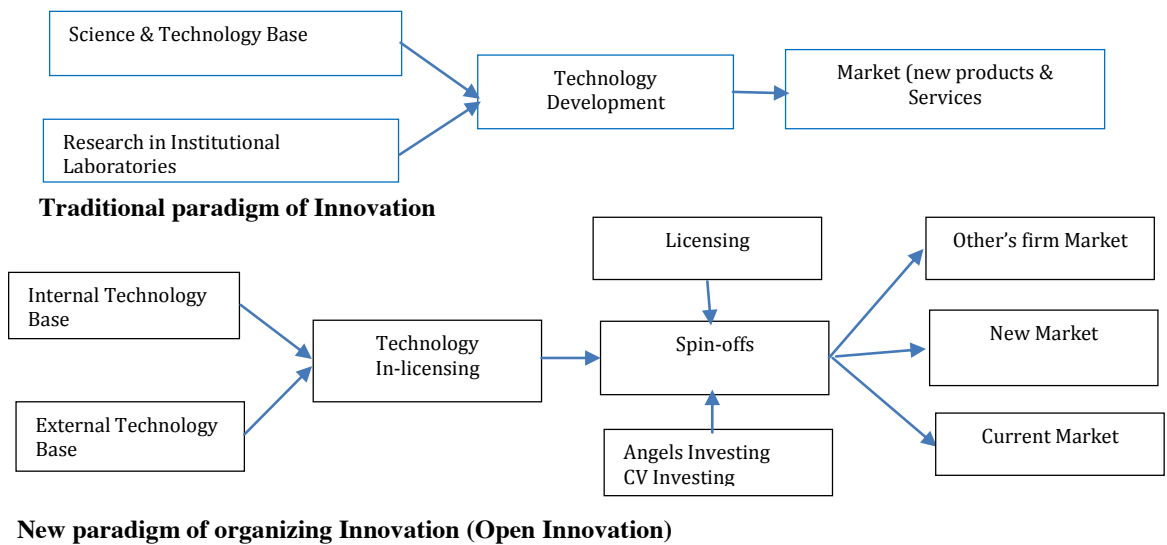


Figure 5: Partnerships and Alliances OI Strategy

Annex 1: Traditional paradigm of innovation vs new paradigm of organizing innovations (OI)



Annex 2: Web 2.0, Enterprise 2.0 and Enterprise 2.1, Web 3.0

The term Enterprise 2.0 is used first by McAfee (2006) in parallel with Web 2.0. McAfee defined Enterprise 2.0 as part of Web 2.0. The difference from Web 2.0 was that Enterprise 2.0 could be used to network internal functions, internal processes and knowledge workers. He characterized it as an emerging technology, low-cost driver of knowledge management with enormous potential of enhancing collaborative work and exchange of information and knowledge among employees. McAfee (2006) outlined the six components of Enterprise 2.0 – Search, Links, Authoring, Tags, Extensions, and Signals – with the acronym SLATES.

Enterprise 2.1 is an application of web 2.0 technologies aiming at networking actors and providers of external information and scientific and technological knowledge sources. To seize the opportunities of these technologies organizations must necessarily reengineer their processes in order to facilitate Web 2.0 technologies utilization to communicate with external knowledge sources.

Web 3.0, or the Semantic web, is a collaborative movement facilitating our ability to collaborate online, by providing the ability to access untapped knowledge and information in online networks. It has enabled a fundamental transformation in open innovation and supply chain management from linear sequential thinking and communication to dynamic communication and collective intelligence involving all innovators around the globe and all levels of the supply chain from extreme upstream to extreme downstream. The Web 3.0 is still in a very early phase of development and use with very little research on the topic, however Web 3.0 is already at work to support and use collective intelligence in open innovation collaboration along the supply chain.

INFORMATION SYSTEMS GOVERNANCE AS A SYSTEM OF RULES: HIERARCHICAL AND HETERARCHICAL IMPLEMENTATION

Research in Progress

Tedjini, Abderrezak, Telecom Business School, Evry, France, abderrezak.tedjini@telecom-em.eu

Morley, Chantal, Telecom Business School, Evry, France, chantal.morley@telecom-em.eu

Soulier, Eddie, University of Technology of Troyes, Troyes, France, eddie.soulier@utt.fr

Abstract

Information Systems Governance (ISG) can be defined as a set of rules allowing executives and stakeholders to determine how they will decide on the Information System management. The first objective of this paper is to propose a set of meta-rules addressing different aspects of ISG, which are instantiated in each company setting. The second objective is to propose two contrasting models of ISG, which instantiate differently the set of rules. Conventional view of ISG includes hierarchical and centralized control with little flexibility to support rapidly changing organizations. Heterarchical forms are more and more frequently observed in ISG practices (agility, transversality, decentralization...). However, if uncontrolled, heterarchy can lead to the emergence of anarchic phenomena, such as instability, increased conflicts, and waste of resources. Approaching ISG through rules implementation can help controlling heterarchical forms. In the first part of our paper, we describe an ISG as a set of the rules, based on Elinor Ostrom's work and her IAD (Institutional Analysis and Development) framework. In a second part, we develop each type of rule first according to a hierarchical view, then to a heterarchical one. Beyond theoretical contribution, the proposed set of rules can help CIOs involved in improving ISG. It can also be used to make an organizational analysis of heterarchical practices of a company's ISG.

Keywords: Heterarchy, system of rules, Information Systems Governance (ISG), hierarchy, IAD framework.

1 Introduction

According to (Reix, 2004), an information system is "*an organized set of resources: hardware, software, personnel, data, procedures ... to acquire, process, store information (data, text, images, sounds, etc.) within and among organizations*". Those resources generally constitute a strategic asset for the company, and they have to be managed. Information System Governance (ISG) relates to how IS resources are managed. It is considered a business imperative for bringing value of information technology. Weill (2004) defines ISG by stressing on decision: ISG is a management process for controlling decisions and the underlying risks. It describes the distribution of IS decision-making rights and responsibilities. Such a management process is usually based on best practices (Jacobson, 2009), and connected to strategic alignment. From a professional point of view (Cigref, 2002, p.11), ISG is "*the combination of steering, which ensure that today's decisions properly prepare for tomorrow, and control*". Cornu-Emieux and al (2009, p.13) consider governance as "*a set of rules and methods to rationally control the information system adequacy with company's objectives and resources*". Thus, we can be consider ISG as a set of rules that allow executives and skateholders to determine how they will decide together, how they will run actions and strategy, and how they will monitor and evaluate them. In practice ISG is often confronted with issues concerning authority, role assignment, information and knowledge sharing, coordination, decisions. The first objective of this paper is to propose a set of meta-rules addressing different aspects of ISG, which are instantiated in each company setting. The second objective is to propose two constrasting models of ISG, which instantiate differently the set of rules. Conventional view of ISG includes hierarchical and centralized control with little flexibility to support rapidly changing organizations. In order to deal with organizational transformations, Chief Information Officers often face complexities and challenges: rapidly changing environment, rigid and slow procedures, growing demand for agile practices... ISG hierarchical model does not resolve those issues because reactivity is not a key concern. Heterarchy could provide an answer to current challenges. This concept refers to polycentrism, partial autonomy, local control, limited centralization, collaboration, transversality, and temporary organizations. A heterarchical system differs from a hierarchical one because it favors decentralization, interrelation and cooperation among members rather than a centralized operation based on a top-down structure. A heterarchy is defined as a polycentric organization, that is to say that more entities (at the same level or at different levels) have a self-regulating, which leads to phenomena of emergence. Some heterarchical forms can already be seen in information systems department practices: project mode, collaborative work¹, cross-functional management, collective decision-making, overlapping entities and interactions with multiple links between members. However, if uncontrolled, heterarchy can lead to the emergence of anarchic phenomena, such as instability, increased conflicts, and waste of resources. Approaching ISG through rules implementation can help controlling heterarchical forms.

In the first part of our paper, we describe an ISG as a set of the rules. In a second part, we develop each type of rule first according to a hierarchical view and then to a heterarchical one.

¹ According to the Gartner Group, collaboration is placed higher in the priorities of French CIOs in 2014 compared to their international counterparts. <http://pro.01net.com/editorial/611860/les-dsi-conscients-des-defis-numeriques-qui-les-attendent-mais-mal-prepares/>

2 INFORMATION SYSTEM GOVERNANCE AS A SYSTEM OF RULES

The social regulation theory has been developed by French sociologist Jean-Daniel Reynaud since the 1970s. It is based on the concept of "rule". Reynaud (1997) tries to understand how rules can sustain social group and collective action. Negotiation and rules are central to social relations. They help understanding current debates and stakes in our society (De Terssac, 2012).

According to Reynaud (1997), "a rule is an organizing principle. It can take the form of an injunction or prohibition to strictly determine a behavior. Most often it is an action guide, a yardstick for judging, and a model for action. It introduces meanings, partitions, and links in the symbolic world."(p.4). One often speaks of "rule of a game", which is a set of principles governing how a game should progress. It sets how we should play while conforming to these rules and requirements. The interdependence of rules constantly interacting, which is part of the concept of a rule system, is of major interest for us. To implement an ISG means setting operating rules. A system of rules includes several types of rules, both autonomous and interacting with others. They apply to "collective action, as the main object of management science" (Teulier and Lorino, 2005).

A major contribution to rules system in connection with organization of collective action is that of Elinor Ostrom's (Ostrom, 1990; Polski and Ostrom 1999; Crawford and Ostrom, 2005). Her theoretical framework represented by the IAD (Institutional Analysis and Development)² model "is used as an analytical framework for the analysis in cases where management is shared. Her analysis focuses on a conceptual unit, called "Action Arena", where actors interact in social situations called "status share" (see Fig.1). In such an arena, participants interact, exchange goods and services, solve problems, dominate or fight in (Ostrom, 2005). Several external factors, called "exogenous variables", affect interactions and outcomes.

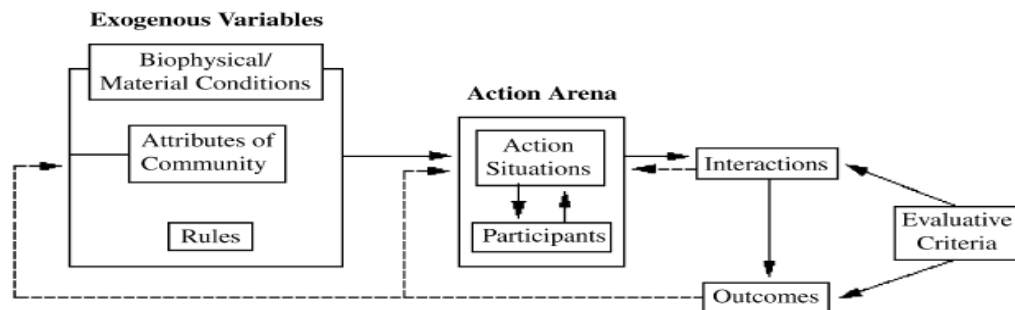


Figure 1. The IAD Framework model (Ostrom, 2005)

Changing the rules which guide the action situations in an arena can help changing the system. Rules are often identified by the actors themselves as the main vector of change. Most often, actors involved in conflict solving consider that establishing new rules is a priority.

Recent research in information systems use the model IAD (Tenenberg, 2008), sometimes concerning agile practices (Montgomery, 2014). Elinor Ostrom has defined a set of rules that we have adapted for

² It is described more explicitly in the work of Ostrom "Rules, Game and Common-Pool Resources" (Ostrom and al. (1994). This model is the IAD collective result of many diverse social scientists who participated in the Workshop in Political Theory and Policy Analysis in the past 25 years (Polski and Ostrom, 1999).

describing an ISG at a meta-level. We have added rules related to coordination and rules related decisions. Our framework includes nine types of rules: R1 to R9. We will give a brief definition of each of them.

2.1 Position rules, R1

Position rules specify which positions, in a given situation of action, may be taken by each category of actors. They help define the actions associated with a given position. To identify the form taken by R1 in a particular context, one should ask: what are the roles of different actors in this situation?

2.2 Boundary rules, R2

Boundary rules specify how actors should be chosen to enter a given role. They help define which actor may participate in the given action situation. To identify the form taken by R2 in a particular context, one should ask for each role: which participants can access the role?

2.3 Authority rules, R3

Authority rules specify the extent of each player's power of control. They provide information on each participant's level of control on action in a particular situation. They specify what a participant in a given position, may, should, or should not do within a decision process. To identify the form taken by R3 in a particular context, one should ask: Which activities are allowed to the participants, and how these activities affect the outcomes of the process? What is each participant's level of control on the process?

2.4 Information and knowledge rules, R4

Information and knowledge rules specify channels of communication between stakeholders, and which information and knowledge may, should, or should not be shared. To identify the form taken by R4 in a particular context, one should ask: which information and knowledge on the situation of the action is available to participants? How information and knowledge are distributed?

2.5 Coordination rules, R5

Coordination rules specify how between actors and units coordinate. To identify the form taken by R5 in a particular context, one should ask: which are methods and mechanisms for coordination? How can collective action be coordinated to achieve tasks? How does one assume coordination?

2.6 Decision rules, R6

Decision rules specify what decisions are assigned to a player in a given position. They provide information on how the action is organized at different nodes of a decision tree. To identify the form taken by R6 in a particular context, one should ask for each situation of action: who decides on what?

2.7 Aggregation rules, R7

Aggregation rules define how the decisions of actors to a node should be connected to intermediate or final results. They refer to generic forms of decision, for example, according to a majority rule or unanimous rule.

2.8 Scope rules, R8

Scope rules define how the actors' behavior affect expected outcomes and by what extent. Will there be room for the initiative? Or should actors comply with the rules? To identify the form taken by R8 in a particular context, one should ask for each situation of action: what autonomy and initiative are left to actors?

2.9 Payoff rules, R9

Payoff rules specify how rewards should be distributed to players, based on their actions and results. To identify the form taken by R9 in a particular context, one should ask for each situation of action: what can players committed to action and results, expect in return?

3 HIERARCHICAL GOVERNANCE VERSUS HETERARCHICAL GOVERNANCE

Minzoni (2010) identifies two contrasting views on governance: hierarchical and heterarchical. The first one is grounded on collective mistrust and desire for control; main answers are frameworks and procedures to mitigate risks. The second one is a kind of utopia, where long-term social cohesion is achieved through multipolar co-controlling. The two governance regimes can be fully apprehended for ISG using our set of rules. For each type of rule, we will give two versions, hierarchical and heterarchical, with examples.

3.1 The position rules, R1

3.1.1 Hierarchical implementation

In a hierarchical variation, each player has an initial position or predefined role on a static chart, at a given hierarchical level, associated with defined responsibilities, and is obviously qualified. For example, a MIS organization can draw a clear hierarchical distinction between "senior developers" and other "junior developers", which are both classified under "software engineering engineer" on the organizational chart.

3.1.2 Heterarchical implementation

In a heterarchical vision, there is a kind of equality between roles, and sometimes roles are redistributed. For example in the "open source" community, there is a principle of egalitarian roles. Software modules are "free" or open source, and each actor can copy, transform, combine them, and then share the results with others, for advancing a collective project (eg linux³, wikispeed⁴).

3.2 The boundary rules, R2

3.2.1 Hierarchical implementation

In a hierarchical variation, the access rule to the role is almost fixed, linked to a stable skill. It is used at one level and can not be exploited optimally at other levels. Therefore, skill is chosen only once, for a fixed role and housed in a single hierarchical level. For example, in a department "purchases" a lawyer is responsible for drafting legal clauses of specifications and can not be assigned to another department. Thus, in a hierarchical view skill is linked to the role and level uniqueness.

3.2.2 Heterarchical implementation

In a heterarchical variation, actors are chosen first of all with regard to their ownership of capacity (capacity for innovation, autonomy, etc.) or skill. This criterion is a decisive element for the

³ Linux is a UNIX-type operating system, its main peculiarity of Linux is to be free software, developed collaboratively and largely volunteer hand by thousands of programmers around the world. <https://aful.org/ressources/presentation/linux>

⁴ WIKISPEED is a United States registered automotive manufacturer and a Washington State, USA, non-profit company with R&D inputs from a global think-tank collaborating using Agile project management and open source licensing (<http://wikispeed.org/>).

establishment of a coherent actor group. Belonging to a hierarchical level intervenes only second. For example, we can find heterarchy in the structure: the division of responsibilities is not purely down (vertical), but can combine geographic, specialized skills, particular local interest.

3.3 The authority rules, R3

3.3.1 Hierarchical implementation

Because they are in charge, leaders often do not try to demonstrate how sound their strategies are. They are not necessarily wrong because the hierarchical principle does not require taking subordinates' opinion. The only thing that is really important is to convince superiors (Dupagne, 2012). The hierarchical operation generally corresponds to a pyramidal distribution of authority, in which lower levels are subordinated to higher levels. However in hierarchical systems, one can observe decision-making process that does not respect the hierarchy.

3.3.2 Heterarchical implementation

In this vision, individualized authority is not popular because the authority is everyone's business and achieved collectively by all. Indeed, the responsibilities of powers are shared in line with the key skills. According to Aime et al., (2014) "*a power heterarchy within teams as a relational system in which the relative power among team members shifts over time as the resources of specific team members become more relevant (and the resources of other members become less relevant) because of changes in the situation or task*" (p.328). For example (Hedlund, 1986) mention a multinational company (MNC multinational corporation), in which R & D center is based in the Netherlands with responsibility for coordinating product development, while a center in Singapore responsible for the Asia market, another center in London is responsible for common purchases across the MNC and the responsibility of the main product is localized in charge of product division in Germany. Unlike a matrix operation, the conflicts must be resolved not by the intervention of a higher authority, but through negotiations.

3.4 The information and knowledge rules, R4

3.4.1 Hierarchical implementation

In hierarchical organizations, information and knowledge are moving to the hierarchy, they are centralized at the top in particular by the decision maker. The withholding of information is shown by Crozier as a source of power (1964). In this vision "the one who holds the information is the one who has the power". Furthermore, knowledge and information are located at each node and are only handled by superiors. If another node needs these resources, it can be accessed only through the hierarchy. This explains the development of individualism spirit: "*I treat information I manage knowledge at my level, to be transferred to my supervisor*". The flow of information⁵ generally follows a hierarchical traffic, rather down, where information spreads from superior towards his subordinates. However, communication networks in the business do not always respect the established chart. Indeed, informal communications have always existed. Spontaneous, they have always been a source of information flow in organizations. For example, information about a possible takeover of the company by a competitor can flow quickly into the business and lead to serious consequences for the mobilization and involvement of employees.

⁵ The flow of information is generally related to the organizational structure of the company.

3.4.2 Heterarchical implementation

Today we are in a predominance of Information and Communication Technologies (social networks, big data, collaborative platform, etc) that allow and demand for information and knowledge sharing at multiple scales. This sharing of information and knowledge useful associated with collective action can develop a transparency strategy that in turn calls confidence to collectively solve related problems. For example, according Ertzscheid "*with Twitter, which dominates at first is the aspect heterarchical deep, flat, information that circulates. There is talk of heterarchy from the time when, in an organization, there is no "higher level" [...] Twitter is, prima facie, a heterarchy: editorialisation no, no "higher level" of information. So Twitter is literally unreadable, because its nature is precisely to reject "setting an order" or prioritizing. Yet Twitter is read*"⁶.

3.5 The coordination rules, R5

3.5.1 Hierarchical implementation

In a hierarchical implementation, coordination is built by directives or standards following a center of command. It is preset in the vertical static chart, i.e., coordination between the superior and his subordinate is translated by contractor and the execution of the order or the submission of obedience. For example Hedlund (1986) speaks of MNC ethnocentric as an organization in which the parent company alone ensures the coordination and control of subsidiaries. It imposes its rules and its control in all subsidiaries. These are usually run by executives from the parent company.

3.5.2 Heterarchical implementation

Coordination is the result of an agreement on knowledge, rules and standards that define the interactions between individuals within a perspective of collective action. The question of what grounds the agreement remains a central issue. According to Ostrom, the problem of collective action is based on a type of coordination that allows the voluntary cooperation of individuals, not their submission to decision-making authority as in the case of the "*hierarchical coordination*" (Hubert and al., 2009). This type of coordination will be ensured by multiple means, and interactions between individuals. Indeed, the heterarchical type of coordination process takes the form of a "*mutual adjustment*"⁷, it represents the weaving of interdependence among individuals, based on egalitarian rules on the principles of collaboration. Individuals assume the coordination of a given collective action. For example, the MNC heterarchical described by Hedlund (1986) is not limited to a single center, and this polycentrism aims at favoring the reactivity and the innovation. New ideas or new products are not created on the initiative and under the control of the parent company, even if the exploitation is then global. Thus, the subsidiary of a country can be the initiator of activities in one area, and then become global coordinator for the development of these activities and at the same time act as a single distributor in other areas.

3.6 The decision rules, R6

3.6.1 Hierarchical implementation

A hierarchical organization operates with reference to a flowchart ranging responsibilities and powers of each entity. In a given node, each actor is responsible for controlling activities, and can make decisions according to the prerogatives that they are assigned to him, and he should report to his immediate superior of the executed results. Here we are in a posture (the donor orders, on the one hand

⁶ Ertzscheid Olivier, "*Twitter: hieratical against the hierarchical*" Affordance.info, ISSN 2260-1856. July 5, 2009.

⁷ This means that coordination between the tasks is ensured through informal communication between individuals (Mintzberg, 1983).

and on the other hand, the execution of the order relying on submission or obedience). In fact, decisions are taken by the supervisor at each level of the structure. The leader, whatever his level, by definition, is responsible for the decision. He is supposed to be the best suited to make decisions, that's why he was appointed. This method of decision-making can be described as "lonely" because "performers" are not consulted. In brief, the leader decides, but is not responsible yet, it is the hierarchy which carries in fine all the responsibility, and the subordinates execute (Delstanche, 2014).

3.6.2 Heterarchical implementation

The concept of delegation of responsibility is very important. The delegation is led by a rational mechanism, "*the principle of subsidiarity*", which guarantees a careful and precise choice of expertise in this area, and encourages operation in a standalone mode at all levels of the organization. For example, in the case studies conducted by (Morley and Bia-Figueiredo, 2013), one project leaders when confronted with a complex specification to treat, requiring it to be distributed among several employees, especially working with a collaborator he calls "*referent*", privileged according to the criterion of trust. Here, the designation of the referent is not related to his position, but rather to trust. Thus, all the actors have side responsibilities themselves to solve simple problems or complex and take appropriate decisions. The empowerment process⁸ expands on multiple levels.

3.7 The aggregation rules, R7

3.7.1 Hierarchical implementation

In a hierarchical vision, sense of decision is always downward. The decision is made by one person. We are close to the image of "*brain of the firm*". Indeed, decisions are linked to hierarchical positions. Decisions are limited to certain categories of people (managers, etc.). Decision making is realized at high levels of the hierarchy and the base executes the tasks without participating in the decision making process. This type of structure includes all mass production companies, such as line production of the automotive industry (Briol, 2008, p.53).

3.7.2 Heterarchical implementation

A heterarchical variation is close to the image of the "*firm as a brain*", where each actor can be considered a manager and make decisions. Heterarchy is seen as a self-organizing management system: "*In self-organizing managerial system, each participant est aussi a manager of this system. Such a system has s'intitule "heterarchy" (hetero: other and archein: to rule), for at Any Time May it be one of your neighbors who is making the decision, you did Reviews another, as the neighbor of others*" (Von Foerster, 1984, p.8). A proper decision mechanism is "*concertation*". A US study (Beierle and Cayford, 2002) based on 239 cases since 1970, gives the following results: dialogue has changed or has helped change final decision in 58% of cases; consultation processes have resolved the conflicts in 61% of cases; they have increased confidence in the agency who led the dialogue in 45% of cases; they increased the understanding of the subject by participating in 77% of cases.

⁸ The process by which a person or a social group acquires control means which enable it to raise awareness, strengthen their potential and transform in the perspective of development, improving living conditions and environment. <http://www.oqlf.gouv.qc.ca/ressources/bibliotheque/officialisation/terminologique/fiches/1298948.html>.

3.8 The scope rules, R8

3.8.1 Hierarchical implementation

In a hierarchical implementation actors have strong propensity to the respect of rules and standards and sometimes without making enough efforts initiative. This develops heteronomy actors in multi levels.

3.8.2 Heterarchical implementation

The expected outcome requires strong reactivity between actors and more agility and independence in problem solving, conflict of explanation and reconciliation of decisions of the place and time. Confidence and increase transparency grows up within the organization. For example, the comparative study between Japanese and US companies made by William Ouchi in the 1980s revealed that Japanese companies success secrets based on the mastery of the concept of heterarchy: "In the early 1980's, the best business in the world were organisms to be found in Japan and, soon I discovered the secret of their success was their mastery of heterarchy". He showed that Japanese companies are characterized by the following: collective decision making, collective responsibility, informal and implicit control procedures, lifetime employment⁹, and evaluations promotions and slow, not specialized careers, global interest. By cons, US companies are characterized by: individual decisions making, individual responsibility, formal and explicit control procedures, short-term employment, rapid assessments and promotions, specialized careers limited interest (Ouchi, 1981).

3.9 The payoff rules, R9

3.9.1 Hierarchical implementation

In a hierarchical implementation, reward systems are classic or obsolete based on extrinsic motivators (carrot and stick).

3.9.2 Heterarchical implementation

The establishment of an incentive system within the group results in optimizing the success criteria, and developing strong participation and commitment within the organization.

⁹ Workers who have the status of "permanent employee" become members of the "community" that constitutes the company that hired them. They have a "job-based membership". This form of employment is very different from that which exists in other countries, namely the "employment-based workstation", where the tasks, the number of hours and the workplace are generally limited. In the Japanese employment system based on membership, employees do not have the right to refuse to perform tasks, work overtime or changing workplace (Hamaguchi, 2013).

Table 1 summarizes the types of rules set out according to the two visions.

| The Rules | Hierarchical implementation | Heterarchical implementation |
|---------------------------------------|--|---|
| position (R1) | Fixed roles; Static Organization Chart. | Egalitarian and dynamic role (roles can be redefined); Action and thinking constantly on multiple levels. |
| Boundary (R2) | Unique access to the role; Competence limited to the role. | Access role multiple times; targeted and variable Competence; Multi level favoured. |
| authority (R3) | Dominant Authority; pyramidal distribution of authority; explicit and centralized control. | Collective and shared authority; dominance and subordination links are reversed and redistributed; tacit and distributed control. |
| information and knowledge (R4) | Withholding information (individualism); Centralized Information and compartmentalized; knowledge retention (individualism); Monopolizing Knowledge. | Information Exchange (collective spirit); Knowledge Sharing (collective spirit); Transparency and trust; Strong corporate culture; Problem-solving using network knowledge. |
| coordination (R5) | Coordination based on directives or standards (control mode); Vertical coordination between superior and subordinate (submission to orders). | Horizontal coordination through mutual adjustment; Strong relationship of interdependence between actors; Focus on common good. |
| decision (R6) | Personal responsibility on top; Decisions is limited to a certain category of actor. | Lateral Responsibility; Delegation of responsibility according to competence; Distributed Decision (several multi-level centres). |
| aggregation (R7) | Centralized Decision on top; Decision is limited to a certain category of actor. | Collaborative Decision, distributed and collegial (by most actors). |
| scope (R8) | Increased heteronomy multi level; Efficacy. | Increased empowerment at multiple levels; Strong responsiveness and agility; Reconciliation of decision making in place and time. |
| payoff (R9) | Classical individual rewards (selfishness). | Collective incentive system. |

Table 1. Summary of the type of rules in hierarchical and heterarchical implementations

4 CONCLUSION

In this paper an Information Systems Governance has been approached as set of mechanisms and operating principles. They have been expressed as a system of rules, based on Elinor Ostrom’s work. Actualized rules shape implemented governance.

The rules may be declined according to a conventional hierarchical vision of ISG, but heterarchical forms can also be introduced here or there. The first contribution is theoretical: beyond the view usually proposed by ISG repositories, IS governance can be expressed at a Meta level. It can be viewed as a political and regulatory arena, made up of rules which express a collective agreement within organizational players on how to control, to steer, to make decision, to assess the company’s information system. Then, from a managerial point of view, a system of ISG rules may help CIOs ask the right questions: how to perform the exchange, collaboration, understanding of roles and responsibilities and the regulation of authority? It is a kind of handbook for CIOs who seek for more agility and decentralization within IS governance, but do not want to merely rely on emerging heterarchical forms.

A heterarchical implementation may lead to "collective leadership"¹⁰ and ultimately harmonize CIOs' managerial practices. The framework helps questioning rules in use and their consistency. Lastly, at a methodological level, it provides guidance when analyzing an organization's ISG, and helps recognize some heterarchical forms. We have just started such an analysis in a large industrial company.

We have planned to extend the scope of investigation in order to consolidate types of rules, and maybe to find additional rules. Future research includes developing an auditing tool, whose functions could be to assess an organization's degree of heterarchy, to identify governance problems, and to suggest possible improvements.

References

- Aime, F., Humphrey, S.E, DeRue, D.S, Jeffrey, B.P, (2014), "The Riddle Of Heterarchy: Power Transitions In Cross-Functional Teams". *The Academy Of Management Journal*, Vol. 57, No 2.
- Beierle, T. C., & Cayford, J. (2002), "Democracy in Practice: Public Participation in Environmental Decisions". *Resources for the Future*.
- Briol, P. (2008), "Ingenierie Des Processus Metiers, de l'élaboration à l'exploitation". Lulu.com.
- Cigref, (2002), Gouvernance du système d'information, problématiques et démarches.
- Cornu-Emieux, R., Poissonnier, H., Verjus, J.-P., Corniou, J.-P., Casademont, S., Tonnelé, A... Vitari, C. (2009), "Stratégie et pilotage des Systèmes d'information". Dunod.
- Crawford, S. E. S., & Ostrom, E. (2005), "A Grammar of Institutions, in E. Ostrom, Understanding Institutional Diversity". Princeton, NJ: Princeton University Press, pp. 137–174. Originally published in *American Political Science Review*, 89 (3) (1995): 582–600., 89(03), 582–600.
- Crozier, M. (1964), "Le phénomène bureaucratique". Réédition Seuil, Coll. Points 1993.
- Delstanche, P. (2014), "Vers un leadership solidaire : La sociocratie : une nouvelle dynamique pour gérer les organisations". Liège: EDI Pro 2.
- De Terssac, G. (2012), La théorie de la régulation sociale : repères introductifs. *Revue Interventions économiques*. Papers in Political Economy, (45).
- Hubert, B., De Turckheim, E., & Messéan, A. (2009), "Concevoir et construire la décision Démarches en agriculture, agroalimentaire et espace rural". Versailles: Editions Quae.
- Dupagne, D. D. (2012), "La revanche du rameur". Michel Lafon.
- Hallépée, D. (2013), "La gouvernance des systèmes d'information", Les écrivains de Fondcombe.
- Hamaguchi, K. (2013), Les problèmes spécifiques du système de l'emploi japonais.
- Hedlund, G. (1986), "The hypermodern MNC - A heterarchy?" *Human Resource Management*, 25(1), 9-35.
- Jacobson, D. D. (2009), Revisiting IT Governance in the Light of Institutional Theory. In 2014 47th *Hawaii International Conference on System Sciences* (Vol. 0, p. 1-9). Los Alamitos, CA, USA: IEEE Computer Society.
- Michael A. West, Joanne Lyubovnikova, Regina Eckert, Jean-Louis Denis *Journal of Organizational Effectiveness: People and Performance* 2014 1:3, 240-260
- Mintzberg, H. (1983), "Structure in fives: designing effective organizations". Prentice-Hall Englewood Cliffs, NJ.

¹⁰ William Ouchi showed that Japan's commitment based on democratic leadership leads to better quality, higher productivity and lower costs while making workers of different levels full partners in the business (Wilken, 2012). In other words, collective leadership will increase rather than decrease the cognitive demand placed on both leaders and team members (Michael and al., 2014). In contrast, the literature on shared leadership and self-management in work teams offers a different perspective on team functioning that suggests that hierarchical authority is not the only power structure in teams (Aime and al., 2014).

- Minzoni, D. A. (2010), Gouvernance ou leadership ? http://www.institut-leadership-bpi.com/images/stories/articles/gouvernance_ou_leadership.pdf.
- Montgomery, A.W. (2014), ScrumBiz: Transforming Culture and Surging Business Team Productivity with Scrum. Présenté à Spring Symposium, Nashville.
- Morley, C., & Figueiredo, M. B. (2013), "Génération Y et pratiques de management des projets SI". Paris; New York: Springer Verlag France.
- Munagorri, R. E. de. (2011), "Introduction générale au droit: Nouvelle édition revue et mise à jour", (3e édition). Flammarion.
- Ostrom, E. (1990), "Governing the Commons: The Evolution of Institutions for Collective Action", Cambridge University Press.
- Ostrom, E. (2005), Understanding Institutional Diversity. Princeton, NJ u.a.: Princeton University Press.
- Ostrom, E. (2010), Institutional Analysis and Development: Elements of the Framework in Historical Perspective. In Historical Developments and Theoretical Approaches in Sociology. In in Encyclopedia of Life Support Systems (EOLSS) (ed. Charles Crothers). Oxford.
- Ostrom, E., & Baechler, L. (2010). Gouvernance des biens communs. De Boeck.
- Ostrom, E., R. Gardner, and J. Walker (1994). Rules, Games, and Common-Pool Resources, Ann Arbor: University of Michigan Press.
- Ouchi, W. (1981). Theory Z: How American Business Can Meet the Japanese Challenge. Addison-Wesley.
- Polski, M. M., & Ostrom, E. (1999), An Institutional Framework for Policy Analysis and Design. © 1999 by authors.
- Reix, R. (2004), "Systèmes d'information et management des organisations" (5e édition). Paris: Vuibert.
- Reynaud, J.-D. (1997), "Les règles du jeu. L'action collective et la régulation sociale" (Troisième édition). Paris: Armand Colin.
- Tedjini, A., Morley, C., & Soulier, E. (2014), Le concept d'hétérarchie : émergence dans différents domaines et intérêt pour la Gouvernance des Systèmes d'Information. Présenté à Actes du 19ème colloque de l'AIM, Aix en Provence, France.
- Tenenberg, J. (2008), "An institutional analysis of software teams". *International Journal of Human-Computer Studies*, 66 (7), 484-494.
- Teulier, R., & Lorino, P. (2005), "Entre connaissance et organisation: l'activité collective". La Découverte.
- Von Foerster, H. (1984), "Principles of Self-Organization — In a Socio-Managerial Context". In P. D. H. Ulrich & D. G. J. B. Probst (Éd.), *Self-Organization and Management of Social Systems* (p. 2-24). Springer Berlin Heidelberg.
- Webb, P., Pollard, C., & Ridley, G. (2006), "Attempting to define IT Governance: Wisdom or Folly?" In Proceedings of the 39th Annual Hawaii International Conference on System Sciences, HICSS '06 (Vol. 8, p. 194a-194a).
- Weill, P. (2004), "Don't Just Lead, Govern: How Top-Performing Firms Govern IT ", *MIS Quarterly Executive*, vol. 8, No. 1, pp.1-17.
- Weinstein, O. (2013), Comment comprendre les « communs » : Elinor Ostrom, la propriété et la nouvelle économie institutionnelle. *Revue de la régulation*.
- Wilken, T. (2012), Heterarchy - The Secret of Japan, Inc., Future Positive.

THE FACES OF BUREAUCRACY: UNDERSTANDING ENTERPRISE SOCIAL MEDIA

Completed Research

Mansour, Osama, University of Borås, Borås, Sweden, osama.mansour@hb.se

Ghazawneh, Ahmad, IT University of Copenhagen, Copenhagen, Denmark, agha@itu.dk

Abstract

The social media are often perceived as tools that support openness and flexible participation among individuals and communities. This might explain the reason why social media have become indispensable for many daily practices in organizations. But how do these organizations appropriate and use these media relative to their formal structures and characteristics is a question in focus for the current paper. Drawing on classical concepts of organizational bureaucracy from Weber (1978) and Eisenstadt (1959), we present a qualitative analysis of empirical data obtained from two large organizations that use a wiki as a collaborative knowledge platform. The results show how the tendency to organize the use of the wiki through introducing structure might create barriers for open and democratic collaboration and knowledge sharing at the workplace. They also show that while a freer approach to using wikis might allow for self-organizing, there is still a possibility for enacting social structures that limit openness and flexibility. As such, the paper contributes novel insights into how social media might be used in bureaucracies and soft bureaucracies.

Keywords: Wikis, Bureaucracy, Organization, Social Media

1 INTRODUCTION

For many organizations, the use of social media has become indispensable for their daily business practices. Organizations continue to incorporate various social media tools and services at the workplace in order to exploit their potential in stimulating participation, engagement, and dynamic exchange of information among their employees (Huang et al. 2013). Many scholars for instance examined the use of social media in enterprise settings for improving knowledge sharing practices (Majchrzak et al. 2013), democratizing the flow of knowledge (Hasan & Pfaff 2007; Wagner 2006), reducing control and flattening organizational hierarchies (Bibbo et al. 2010) and accelerating innovation and the development of new products and services (Zwass 2010).

Social media are defined as a group of internet-based applications that build on the ideological and technological foundations of web 2.0 (Kaplan & Haenlein 2010). The wiki technology, which is studied in this paper, represents one kind of social media technologies that is often used by organizations to improve their collaborative and knowledge sharing practices. It is described as a simple technology that allows open, organic, collaborative and incremental development of content (Cunningham 2004). Since wikis are often used by communities to support their collaborative knowledge practices (Yates et al. 2010; Bibbo et al. 2010; Mansour et al., 2011) they have been described as 'next generation knowledge management systems' (Pfaff & Hasan 2007). In general, social media are often seen as a new class of technologies that have the potential to alter organizational processes in fundamental ways (Treem & Leonardi 2012; Huang et al. 2013).

While we believe and to a certain extent agree that social media may have brought new affordances for organizations (Mansour et al., 2013; Treem & Leonardi, 2012;), in this paper we scrutinize the potential of social media and its impact on organizational practices. The question of how social media can affect organizations and how they may transform organizational practices and structures may not be clear yet (Majchrzak et al. 2013; Treem & Leonardi 2012; Saldanha & Krishnan 2013). For instance, Aral et al. (2013) noted that research on social media has rapidly increased in some areas whereas others are nascent or non-existent. With respect to organizational wikis, only little knowledge is known so far about how they are used behind the firewalls of organizations (Mansour et al., 2013; Martine et al. 2013; Danis & Singer 2008; Grudin & Poole 2010; Holtzblatt et al. 2010). So current studies do not contribute enough understandings of why wikis succeed or fail, how they work and what possibilities they may afford in organizations (Martine et al. 2013).

Fortunately, a growing amount of research has recently been published that point to problems, dual impacts, or contradictory influences of using wikis in organizations. Grudin and Poole (2010) suggested that problems may arise when using wikis in organizations because of possible incompatibilities between the wiki bottom-up character and hierarchical top-down structures of organizations. Further emerging perspectives in this area also centre on contradictory and dual influences of using wikis in organizations (e.g., Hildebrand et al. 2013; Mansour et al., 2011; Majchrzak et al. 2013).

In light of these arguments, this paper offers a qualitative empirical analysis of wiki use practices in two large, multinational organizations where the wiki technology is used as a knowledge management platform. It aims at examining and understanding the ways by which wikis are used within formal organizational structures. The central question in the paper is *how do organizations appropriate and use wikis relative to their formal structures and characteristics*. In order to answer our research question, the paper draws on key theoretical concepts from the literature on organizational bureaucracy (e.g., Eisenstadt, 1959; Weber, 1978). The remainder of the paper proceeds as follows: the next section provides an overview of the literature on the use of wikis in organizations, and IT use in relation to organizational bureaucracy. Then, a discussion of key theoretical concepts on bureaucracy is presented in section three. The empirical settings and the data collection and analysis processes are described in section four. In section five, an illustration of the empirical findings is

provided. Finally, the paper ends with a discussion of the main arguments on the bureaucratic use of social media and an outline of key conclusions.

2 RELATED LITERATURE

2.1 Wiki Use in Organizations

A wiki is a simple technology that consists of a set of dynamic, interrelated web pages where people can continually create and edit content (Happel & Treitz 2008; Yates et al. 2010; Hester & Scott 2008). One key property of wikis is collaborative authoring and publishing of content (Majchrzak et al. 2013; Hasan & Pfaff 2007). Wiki content is often produced by many people and is organized by topic and subtopic compared to the chronological organization of content in discussion forums.

Majchrzak et al. (2006) conducted one of the earliest studies on organizational use of wikis. They suggested that wikis might produce three key benefits for organizations: enhanced reputation, make work easier, and help an organization improve its processes. Basically, organizations often seek to use wikis to facilitate the creation and exchange of knowledge among communities (Bibbo et al., 2010). Community members collaborate together to develop evolving, continuously-updated content on the wiki taking advantage of its editability (Treem & Leonardi, 2012). This is enabled by 'knowledge shaping' which is an activity that involves rewriting, integrating, and restructuring content (Yates et al. 2010). Shaping wiki content, Majchrzak et al. (2013) elaborated, is a synthesis and organizing activity in that it involves publicly modifying and reorganizing content, removing redundancies and inconsistencies, and making content more meaningful, usable and maintainable.

By and large, recent studies show that wikis can be used in various forms and produce different consequences for organizations. Wikis can potentially promote an open and democratic organizational attitude to managing knowledge (Hasan and Pfaff 2007). Bibbo et al. (2010) also discussed that wikis may help business organizations in fostering a culture of collaboration which encourages employees to willingly create and share knowledge with each other. Further, wikis can also be used to increase transparency in organizations (Danis & Singer 2008). In a recent article, Majchrzak et al. (2013) explained that wiki users have the possibility to exhibit their knowledge breadth by exposing themselves to more viewpoints or contributing to multiple wikis in other disciplines.

But despite these opportunities, organizations are often faced with problems and challenges when applying wikis at the workplace. Holtzblatt (2010), for instance, identified a number of factors that impede wiki use in enterprise settings such as reluctance to share, extra cost for sharing, information sensitivity, unwillingness to share unfinished work, sensitivities to the openness of information, work practices, and many other factors. In a similar vein, Grudin and Poole (2010) discussed success factors for the sustainability of enterprise wikis. They argued that, while organizations apply wikis to provide a dynamic repository of knowledge, these initiatives are often faced with challenges since management visions do not match the benefits of successful wikis. In addition, Mansour et al. (2013) identified a set of organizational wiki affordances that includes commenting, viewability, accessibility, and validation. These affordances suggest that wikis can be used for limited accessibility where content is subject to reviews by the management which is a barrier for flexible and transparent knowledge sharing and collaboration.

2.2 IT Use and Organizational Bureaucracy

The bureaucratic, rational view of organization is the most dominant in Information Systems (Checkland & Holwell 1998). This view upholds the hierarchical structure which makes for "*the original thread from which the fabric of organization was woven*" (Zammuto et al. 2007 p. 750). The hierarchical structure of organizations then defines authority relationships, specifies information flows, and provides a medium for the control and coordination of activities in organizations (Zammuto et al.

2007). The technical superiority of this mode of bureaucratic, hierarchic organizing is what makes it more dominant than any other forms of organization (Weber 1978; Mintzberg & McHugh 1985).

A typical hierarchical organization is often comprised of people and resources which are combined in a suitable structure that is used to take decisions to achieve defined goals and purposes. This simple understanding of organization is discussed by Checkland and Holwell (1998) and highlights the basic elements of an organization: people, resources, structure, and goals. They argued that much of IS literature applies this model of organization that represents the hard, functional strand influenced by the works of Herbert Simon. In this view, structure is central to facilitating organizational decision-making which is seen to be the fundamental managerial activity (Zwass 1992; Checkland & Holwell 1998). The dominance of this model in IS literature can be referred to the lack of an alternative form which advocates an interpretive, soft view of organization that may challenge the rational, goal-seeking mechanistic view. Checkland and Holwell (1998) even argued that interpretive IS researchers did not offer a model of organization that builds on their soft approaches. One reason for this is that the bureaucratic form of organization tends to be more suited to the fundamental nature of computer systems in that organizations are programmed to perform predefined, formal tasks (Dahlbolm & Mathiassen 1993).

In addition, the role of IT in bureaucratic organizations is discussed by Zammuto et al. (2007). They explained that in the early days IT was used by organizations to automate existing organizational practices, thus enforcing the already existing hierarchy – something they called automated plumbing. IT was therefore taken for granted and deemed negligible. But nowadays, advancements in IT capabilities may be supplanting organizational hierarchies as they offer more flexible control and coordination of activities in organizations. Surely, advanced technologies introduce novel affordances and induce organizational change but whether and how these advancements can fundamentally change organizing is still an issue subject for debate and scholarship.

3 THEORETICAL BASIS

The concept of bureaucracy has been thoroughly researched and investigated by researchers in various research settings (Adler and Borys, 1996; Casey, 2004; Eisenstadt, 1959; Kallinikos, 2004; Vaast, 2007). There are two general viewpoints on bureaucracy. The first view defines bureaucracy as a tool that is developed to efficiently implement particular goal(s). The second view looks at bureaucracy from a power perspective and defines it as an instrument that is used to exercise control over people (Eisenstadt, 1959). Researchers have assessed these two viewpoints, which are not entirely contradictory, from positive and negative assessments (Adler and Borys, 1996): The negative assessment (cf. Walton, 1985) sees bureaucracy as relying on a traditional command and control model (Adler and Borys, 1996). The positive assessment highlights the technical function of bureaucracy and sees it as a coordination mechanism that encourages humans to accomplish their assigned tasks successfully (Deeming, 1986).

Bureaucracy is an organizational topic by nature. Understanding the development of bureaucratic organizations is among the most researched topics by scholars. According to Eisenstadt (1959) the development of bureaucracy relies on various factors: (1) the differentiation between roles (e.g., economic, political, etc.), (2) the differentiation between membership and achievement, (3) the differentiation between functional groups (e.g., board of directors vs. field engineers), (4) the differentiation of the definition of a community, and (5) the differentiation in the social structure. These force different kinds of institutional units inside organizations to compete for resources and support to achieve their goals while practicing, facing and dealing with bureaucracy overtime.

Studying bureaucracy in organizational contexts is tightly connected with authority (Casey, 2004). This has been explained by the theory of bureaucracy developed by Weber (1978) where the rulers (e.g., managers, business owners) believe that they have the right to exercise authority of their own will, and the ruled (e.g., employees, workers) obey the demands is unquestionable duty. Weber's

(1978) efforts to explain this led to a “rational-legal” of bureaucracy that is seen as the “legitimate authority within bureaucratic organizations” (Casey, 2004, p.62). The rational-legal rules separate between the rules and humans administrating them within organizations. This differentiates modern organizations from other types of organizations of its “rationalization of authority and legality” (Casey, 2004, p.62). This rationalization maintains the domination of bureaucratic administration through knowledge (Weber, 1978). It exemplifies the practices of managers encouraging and facilitating the participation of their employees in workshops, seminars and training programmes to gain knowledge that aids decision makers (Casey, 2004).

In his study Vaast (2007) reported how knowledge management systems directly impact a bureaucratic environment. The study shows how online practices impact bureaucratic offline practices. The emergent continuity between both practices lead to the practice of soft bureaucracy. It is identified by Courpasson (2000) as “an organization where processes of flexibility and decentralization co-exist with more rigid constraints and structures of domination” (p. 157). Top management used to promote the concept of soft bureaucracy to maintain the self-organizing for their knowledge workers while streamline the organizational processes (Hodgson 2004; Karreman and Alvesson 2004). Vaast (2007) emphasized that the trend of soft bureaucracy that is a result of online and offline practices should be identified throughout the organization by examining the conditions and processes in which bureaucracies are developed.

4 EMPIRICAL METHOD

4.1 Enterprise Settings: CCC and IBM

The current study took place at two large, multinational organizations: CCC and IBM. Each uses the wiki technology as a platform for knowledge collaboration and sharing. The first organizational setting is CCC. CCC is an acronym which stands for Consolidated Contractors Company. It is one of the largest construction companies in the world and is specialized in offering a myriad of construction, contracting, and engineering services. It was founded in 1952 and is currently ranked number 13th in the top international contractors list by the Engineering News Record magazine. The company has over 170000 employees in more than 100 countries. It is worth noting that CCC has increased its manpower from 25000 in 2003 up to 170000 nowadays. The company was thus confronted with various challenges due to this exponential leap in the numbers of employees including lack of and inconsistent communication across projects, high levels of decentralization, complex knowledge transfer and many other challenges. The company then decided to establish a Knowledge Management (KM) department responsible for setting a strategy for a corporate knowledge network in 2007.

The KM department was first tasked with finding a collaborative platform that would help the company in building a ‘knowledge rich culture’ and allow them to leverage expertise and knowledge. The department decided to use a wiki as a collaborative platform for knowledge collaboration and sharing practices. The wiki was officially launched in 2008. It was mainly used by what the company call ‘communities of practice’. These communities represent various professional groups who use the wiki as a central platform to collaborate and share knowledge obtained through projects. They often share, for instance, method statements that describe specific and formal techniques and procedures used in real-life projects. These communities are structured, that is, each community has a specific structure defining roles and relationships among the community members. So there are community leaders, community captains, subject matter experts and other regular members. The wiki is only accessible by community members. Anyone interested in a particular professional area is expected to send a request to the KM department that decides whether to grant accessibility or not. The wiki is therefore fully controlled by the KM department and can only be accessible through an internal, secure CCC network. Further, the wiki is not ‘open’ in the sense that it is not possible for anyone to create or edit articles - that mostly depends on the defined role of each community member. When a member is allowed to share content on the wiki, it is often the case that a senior community member would

review it and decide whether to provide approval for publishing or not. The wiki is currently used by 11 communities and the company aims for up to 50 communities.

The second organizational setting is IBM. It is also a large, multinational company and often referred to as the largest software company in the world. IBM is mainly specialized in developing software and hardware products and also offering consulting and hosting services. One important product is IBM Connections. It is a universal social platform that includes a myriad of social collaboration tools such as wikis, blogs, status updates, file sharing, ideation, user profiles, community spaces, and other social tools. Connections, as people inside of IBM call the platform, was first released by Lotus Software which was owned by IBM in 2007. The focus in this study was on the wiki part of Connections. Wikis at IBM can be created by anyone and can be used for various purposes. So there is no central control of the use of wikis. Any employee within IBM can, for instance, create a wiki for a private project team, public community, personal publishing tool, and many other purposes. Access to these wikis is determined by their purposes. So a public community can be virtually accessed by the 400000 employees who work at IBM. But a wiki created to serve as a platform for a private project team can only be accessed by the team members. People who create wikis can control them. Anyone who creates a wiki can determine how public or private a wiki can be, who can access it and what possibilities a wiki can afford (e.g., editing, commenting, etc.). There are now thousands of wikis at IBM which are used for various professional and social purposes.

4.2 Choice of Method and Data Collection

The qualitative interview method was the primary vehicle for collecting empirical data at both CCC and IBM. The choice of using the interview method was primarily motivated by our interest in understanding individuals' use practices of the wiki technology in enterprise settings. This involved focusing on peoples' experiences, actions, thoughts and many other forms of social practices. This suggests that in order to understand the use of technology there is a need for a method that can help in obtaining the meanings of peoples' experiences and also engaging with them directly. The interview is the most suitable method to achieve such purposes (Schultze & Avital 2010; Kvale 1996). Also, as a tool for collecting qualitative empirical data (Schultze & Avital 2010), the interview method is considered a powerful research tool (Kvale 1996) that can help in developing insights into human thought and action (Klein & Myers 1999).

| Data Sources | Description |
|--------------|---|
| Interviews | - The total number of interviews: 38 Company: 27 interviews at CCC over two phases, 11 interviews at IBM. Type: 15 interviews over phone, 11 over Skype and 12 Face-to-Face. Country: 4 interviews in USA, 17 in Greece, 2 in UK, 4 in Denmark, 1 in Qatar, 1 in Spain, 2 in KSA, 1 in Australia, 3 in Oman, 1 in Kazakhstan and 2 in UAE. Interviewee's role: 6 project managers, 9 engineers, 2 social media evangelists, 13 managers, 3 IT professionals, 1 HR specialist, 2 technical professionals, 1 sales professional and 1 learning intelligence leader. |
| Documents | Official documents: KM presentations, stats on the wiki, and monthly newsletters. |
| Field Work | Community meetings: 2 field visits and live observation of the wiki in action. |

Table 1. Summary of data sources at CCC and IBM.

The total number of interviews at the two companies was 38. A summary of these interviews is shown in Table 1 above. 27 interviews took place at CCC over two phases of data collection. 15 interviews were completed in the first phase and 12 in the second. The second phase of data collection at CCC included participants interviewed in the first phase so a number of participants were interviewed twice in different periods of time. In parallel to the second phase at CCC, 11 interviews were also performed

at IBM. It should be noted that the number of interviews at CCC is higher than in IBM because several of CCC participants were willing to be interviewed twice. Also, the data collection was part of a larger project which was time-limited. The collection of data at IBM started later within the project time which created limitations for further investigation at the company.

There was diversity in the participants' seniority, levels of expertise, cultural backgrounds, and wiki use experience that can be seen in Table 1 above. Interviews in both CCC and IBM were performed through various channels such as face-to-face discussions, Skype, and telephone communications. The reason for this is the geographical distribution of the participants over a number of countries. The average interviewing time was 50 minutes and all interviews were recorded using an audio recorder, transcribed, and later verified by the participants.

While interview data represent the backbone of our analysis, other sources of data mainly official organizational documents and field visits were used to provide further support for our empirical data set. It is important to note that obtaining documents and visiting the field was only possible at CCC. In general, documents allowed us to get insights into actual wiki uses in terms of number of contributions, interaction frequency, and so on. Field visits – two visits took place at CCC headquarters in Greece and Abu Dhabi to attend community meetings – were useful to obtain practical insights into how community members use the wiki at CCC. During the meeting, the wiki interface was projected into a screen and community members discussed different issues and put it live on the wiki. It has overall given us a sense of power relations among members and the adaptation of wiki functionalities to meet the community's needs.

4.3 Data Analysis

The analysis of empirical data follows the hermeneutic tradition of analysis. That is to say we sought an explication of the meaning of the qualitative text obtained through interviews (Cole & Avison, 2007). A number of key theoretical concepts of bureaucracy were applied during the data analysis to support theoretical development based on the meanings implied in key empirical instances. These concepts include bureaucracy as instrument, bureaucracy as a coordination mechanism, roles, membership and achievement, functional groups, community, social structure (Eisenstadt, 1959), and the different types of domination (Weber, 1978).

The application of these concepts was enabled by the hermeneutical circle (Cole & Avison, 2007). It is described as a mechanism for analysing textual data in an iterative, spiral manner (Klein & Myers, 1999; Cole & Avison, 2007). This was applied in our analysis through examining the meanings of different parts in the text and then establishing meaningful relationships across them. Two key interpretive processes were used as means to explicate the meanings in the text: *understanding* and *explanation* (Cole & Avison, 2007). The actual data analysis was therefore performed with emphasis on developing certain interpretive understandings of text in light of chosen theoretical concepts and then explaining these meanings for the development of our theoretical arguments.

Each individual interview transcript was examined by both authors separately but using the same analytical processes. This involved first reading and understanding the text in each interview transcript and then explaining the data based on theoretical concepts. We used a simple table (cf. Miles & Huberman, 1994) comprised of three main columns. The columns include the theoretical concepts, empirical quotes, and a brief explanation of how the concept is related to the quote.

Initially, understanding the data in terms of identifying key empirical instances and linking these to theoretical concepts was the main focus. This was an interpretive effort concerned with making sense of the data in order to establish meaningful links and connections to theoretical concepts. The aim from this interpretative understanding was achieved by documenting several key empirical instances in relation to certain theoretical concepts in two separate tables developed by both authors of the paper. The tables were later compared in order to summarize dominant instances and concepts in one shared table to be further examined.

Then, for further examination and interpretation of the data, we started to develop explanations of the links that we found between the data and the theoretical concepts in the previous understanding step. So focus was placed on elaborating these linkages in an attempt to do interpretive reflection and reconstruction of the data relative to theoretical concepts. Cole & Avison (2007) described this as “*shared meaning is interpreted anew*” (p. 825). We re-examined the data in the summarized table with emphasis on eliciting potentially new meanings from the data and developing theoretical understandings that satisfy the core focus of each concept. An explanation was thus developed for each of concept and its relevant empirical quotes. Hence, the development of these explanations aimed at providing a basis for discussing our theoretical arguments primarily in relation to identifying different forms of bureaucratic practices in the context of using a wiki in enterprise settings.

5 EMPIRICAL FINDINGS

5.1 Wiki Use at CCC

The use of the wiki at CCC is essentially bureaucratic. There are many dimensions of the bureaucratic use of the wiki in this organization. One key dimension is related to the structure of communities of practice that use the wiki. As stated in the description of the organizational setting above, the wiki at CCC is mainly used by professional communities and anyone who wishes to use the wiki needs to be a member in one or more of these communities. The bureaucracy here can be understood in terms of the hierarchical structures and authority relations that determine the roles and tasks of each member as well as the flow of knowledge within the community. In fact the structure of these communities reflects in many ways the dominant bureaucracy of the organization. That is, each community has some sort of ‘top management’ that determines the subjects to be shared and discussed on the wiki, reviews any content shared by members and makes decisions about who can join in the community. A Mechanical Manager, for instance, commented on how community leaders and captains are responsible for verifying and reviewing content on the wiki:

“We have many procedures in the pre-commissioning community. We have four captains to approve these procedures.”

Community leaders and captains also meet offline to discuss different issues related to the community and their wiki collaboration. These meetings are essential to the growth and continuity of each community because senior members meet together to implement plans that they agree on and decide action items to move with the community forward. Mostly, these are formal meetings led by senior community members with a ‘modest’ participation of junior employees. One of the Control Quality Managers explained the role of these meetings:

“it is not only exchanging ideas online. We have meetings ... people of these communities meet and discuss things, and the thing is that this is some kind of filtering and coming up with better ideas. Also it is about coming up with consensus and agreement on these ideas. So it’s not only writing and reading.”

Equally importantly, content shared across these communities is not openly accessible by everyone. The issue of accessibility to the wiki is actually an important dimension that reflects the enriched bureaucratic thinking in the organization. In many cases, people, especially junior employees, had difficulties to access content on the wiki. Usually, controlled accessibility to the wiki is a preference for senior managers. One of the Plant Group Managers strongly explained his opinion:

“the wiki is not Facebook where it is completely open ... You only invite certain number of company employees to share their knowledge.”

The manager here makes a reference to Facebook to emphasize that the wiki should not be open to everyone like in social networks and that it should only be used for professional purposes. This tendency to control access to the wiki resulted in problems for junior and less experienced employees.

Salma, an administrator, explained her concerns with accessibility and the need to submit a formal request to get access to the wiki:

“...you need to send a request and mostly they will approve it. But some people think that is limited for only specific people. And so maybe this perception influenced my opinion.”

In fact, CCC treats content shared on the wiki very seriously in that it considers any content to be some kind of formal professional procedures or method statements that represent actual knowledge and experience. This is what gives senior community members privileges in using the wiki and controlling shared content. A Control Project Manager explained:

“Although we share ideas we may reach disagreement...Then some more senior people should interfere ... so it's not only about sharing the knowledge. At the end of the day, when it comes to real execution of the job, somebody must have a say and say yes I agree. This is the way to do it.”

In addition, the bureaucratic use of the wiki at CCC takes many forms as can be seen in the empirical instances above. Each of these instances reflects some dimensions of bureaucracy such as control, hierarchy and formal relationships. Such factors drive a bureaucratic use of the wiki in the sense that it is not used for informal, flexible and open collaboration. Interestingly, these factors are essentially rooted in the formal top-down structure of the organization.

5.2 Wiki Use at IBM

While the use of the wiki at IBM may seem less bureaucratic compared to CCC, still there are some intriguing forms of bureaucracy that affect the use of the wiki. Bureaucracy at IBM does not take the typical form of control and hierarchy. It is mainly found in practice as people attempt to use the wiki for various collaborative purposes. That is to say, people are not entirely constrained by the organizational hierarchical structure or any other established organizational forms (e.g., formal rules, standard procedures). Instead, people tend to enact and engage in bureaucratic practices that often limit their opportunities for using and exploiting the wiki. These practices are not necessarily formal but people create some kind of ‘informal boundaries’ that exhibit some kind of bureaucracy. Many groups and communities that use wikis for a variety of purposes often develop informal rules and guidelines to govern and manage their use of the wiki. These then become standards for how each community member is expected to use the wiki. A User Experience Professional described how his team develops internal agreements to assign each member of the team a specific role:

“The whole wiki is open for everybody. But we just have an agreement, okay, here is the master writer for this, and Sally is the master writer for this one, and Bob is the master writer for this one, and everybody else just comments.”

Similarly, software developers also develop internal project guidelines within their teams to manage their collaboration using the wiki. One Expert Software Developer said:

“We created a wiki where we agreed about the guideline and then we just worked from there...”

Even on a personal level, there are people who can create their own wikis and use them as a way to publish personal thoughts and experiences; something that is reminiscent of a blog rather than a wiki. A User Experience Professional said:

“...my experience is that some people are using the wiki technology as just a simple way to publish things so instead of using a blog or a word document they're actually using wikis not in the Wikipedia sense that says my goal is to create a page and let everybody else to make it better...”

Further, sometimes wiki users bound themselves by their professional roles when it comes to engaging with others to develop wiki content. Here, for instance, a Client Technical Professional talks about how she sees her own role relative to others and how it affects her possible contribution into the wiki:

“... if it is a product documentation wiki I wouldn't go in directly in that document... that is because I am a user and the other ones are the kind of producers ... I mean the producers that have this as their job responsibilities...”

These practices suggest that wiki users at IBM may tend to indulge in bureaucratic wiki uses by creating informal rules and roles as well as developing certain perceptions or interpretive schemes of how the wiki works that often limit them from exploiting the potential of the wiki. It is bureaucracy because the aim is to control and govern the use of the wiki and to make wiki collaboration more formal and bound by professional rules and norms at the workplace. This is especially true when wiki users come from different professional backgrounds and levels of expertise. In IBM, people are very sensitive to these issues in relation to their use of the wiki. A Learning Intelligence Leader explained his view, for instance, concerning others editing his content on the wiki which emphasizes his attitude as an expert, he said:

“I am a quite senior resource within our team. I would be very surprised to see other team members editing my manager's post or even editing my post.”

Another similar view shared by an Expert Software Developer when he was asked for his opinion of someone else editing his content:

“I would find that this person is breaking a social habit. Without contacting me first and putting a comment or anything that would be a bit weird.”

The view held by the software developer implies an important dimension of bureaucracy. Despite the lack of formal rules in the wiki environment, still there are informal social practices that require wiki users to behave in certain ways. In this case, the expert software developer expects others to contact him if they wish to contribute anything to his content on the wiki. In other cases, people at IBM do some kind of self-organizing in that they only contribute and collaborate with people of similar backgrounds and expertise. This was explained by an Information developer, she said:

“...when I am collaborating with a software developer I tend to comment. But when I am collaborating with other information developers I tend to edit.”

Another aspect, which suggests the tendency for formal wiki collaboration at IBM, can be found in the kind of content shared on the wiki. A Sales Enablement Professional explained:

“We're not personally invested in the wikis and the kinds of ideas that we share on the wiki, it is never an opinion or it is never a discussion it is always facts.”

Finally, a Project Manager who has an assignment to setup wikis and monitor them reflected her general opinion about sharing content on the wiki inside the company, she said:

“So it is more like their knowledge sharing, more than it is actually people sharing knowledge.”

6 DISCUSSION

The application of social media in enterprise settings is often driven by the interest in supporting dynamic and flexible collaboration and knowledge sharing. Our focus in this paper is on examining and highlighting how wikis are appropriated and used within the formal structures of an organization.

As suggested by our empirical findings from CCC, one key form of the bureaucratic use of the wiki is the structure of communities of practice. These communities are the backbone of the wiki platform; they provide a medium for community members to collaborate and share knowledge. However, the structure of these communities reflects the two views of the role of bureaucracy by Eisenstadt (1959). One the one had, structured communities is a tool to achieve the purpose of creating and sharing knowledge collaboratively using the wiki. On the other, it is a tool to exercise control not necessarily over people but over content generated on the wiki.

This can also be seen in the data in relation to formal review procedures and offline meetings among senior community members who review and approve content. Here the differentiation in roles and functional groups (e.g., junior employees, senior employees) (Eisenstadt, 1959) is what makes for the legitimate authority discussed by Weber (1978). Authority and power relations are fundamental in CCC. For instance, anyone who wants to be a community member is required to submit a request to their managers as well as the knowledge management department. Content creation and sharing is also a subject of rationalization of authority (Casey, 2004; Weber, 1978). There is a tendency by managers and seniors experts to show dominance over content which results in barriers for other junior members of the community. Editing, contributing and manipulating wiki content maybe difficult because dominant community members may restrict wiki affordances (Treem & Leonardi, 2012; Mansour et al., 2013) for others.

The use of the wiki at IBM reflects a soft bureaucracy (Vaast, 2007; Courpasson, 2000). The data obtained from this company show that it has a culture of openness and flexibility that encourages employees to use the wiki for dynamic collaboration and knowledge sharing practices. These practices however are still bound by existing differentiations in roles, functional groups, and social structures (Eisenstadt, 1959). The different elements discussed by Eisenstadt make for a soft bureaucracy in that wiki users are somehow limited by their professional roles and affiliations as suggested by the data. Employees within the same professional group and who have similar roles (e.g., software developers) would find it easier to share with each other compared to sharing with other professional groups.

Perhaps what might be an explicit element of soft bureaucracy at IBM is self-organizing (Hodgson 2004; Karreman and Alvesson 2004). Wiki users tend to self-organize by enacting both informal and formal rules to organize and govern the use of the wiki. For instance, informal agreements, guidelines or habits are often discussed among wiki users about who to contribute and what to contribute. In other instances, there are people assigned with 'formal tasks' to edit content or to publish content onto the wiki. The data suggests that such rules serve as functions or drivers for soft bureaucracy. That is, they help as coordination mechanisms among wiki users to support them in achieving their purposes from collaborative practices (Deeming, 1986). The community structures at CCC are relevant to such formalization of tasks through which certain wiki users become some sort of wiki managers (e.g., community captains, subject matter experts).

Self-organizing in a soft bureaucracy was observed to lead to certain forms of organizing that are not necessarily conducive to flexible collaboration and sharing with wikis. Employees at IBM for instance have the freedom to setup wikis and use them as a blog for publishing and sharing personal opinions. This may transform wikis into what might seem like a restricted environment and wiki ownership becomes a tool to control content. That is not to say that this kind of organizing is a form of rational bureaucracy (Weber, 1978) but it may lead to the development of guarded wiki silos. Another relevant form of organizing is probably related to differentiations discussed by Eisenstadt (1959). While the data from IBM suggests that bureaucracy primarily functions as a coordination mechanism for wiki users, self-organizing was also a mean for them to account for their professional roles. It was observed that IBM wiki users sometimes discussed their professional roles and expertise in relation for instance to contributing and editing content onto the wiki. That was neither an exercise of control nor an effort to legitimize an authority. But it can be argued that the levels of flexibility and openness might be constrained when expert wiki users demand that they must be consulted sometimes about the content they contributed into the wiki. Others might therefore be discouraged or tend to be limited by what might seem to them a sort of control and domination over content.

7 CONCLUSIONS

The current study is focused on examining the ways by which social media are used within the formal boundaries of an organization. It aimed at highlighting how bureaucratic characteristics of organizations such as authority relations, domination, control, differentiation in roles and functional

groups, and others affect how wikis are appropriated and used at the workplace. The study of two organizations with different structural setup led to various insights into our focus on social media use and organizational bureaucracy. It was shown that community structures that aimed at providing coordination mechanisms for the organization of community activities on the wiki were a cause for command and control. The reason for this can be due to the potentially flexible and open ways of content creation and generation on the wiki. The management in each community was concerned about the validity and accuracy of content, expertise levels of community members, and also their tendency to impose control. Community structures were thus seen as a backbone for a command and control hierarchy in the wiki environment. Relevant social structures existed in the other organization but for completely different reasons. The flexibility and openness of using the wiki allowed people to self-organize and gave them the freedom to be in control of what they create and share. This fostered and encouraged people to openly collaborate and share knowledge with other. But self-organization was also a mean for others to limit openness and flexibility by giving individuals the power to control how they use the wiki and how to let others approach their content. A soft bureaucracy can thus be enacted due to discontinuity between control and flexibility or openness and protection. There are many faces of bureaucracy that can be enacted and the degree to which this bureaucracy affects technology and its use in the workplace will always vary relative to unique organizational characteristics, hence resulting in different consequences of technology use.

References

- Alder, P., and Borys, B. (1996). Two types of bureaucracy: enabling and coercive. *Administrative Science Quarterly* 41(1), 61 – 89.
- Bibbo, D., Sprehe, E., Michelich, J. and Lee, Y. (2010). Employing wikis as a collaborative information repository in a media and entertainment company: The NBC universal case. In proceedings of the *31st International Conference on Information Systems*, St. Luis, USA
- Checkland, P. and Holwell, S. (1998). *Information, Systems, and Information Systems: Making sense of the field*. John Wiley & Sons, West Sussex.
- Casey, C. (2004). Bureaucracy re-enchanted? Sprit, experts, and authority in organizations. *Organization* 11(1), 59 – 79.
- Cole, M. and Avison, D. (2007). The potential of hermeneutics in Information Systems research. *European Journal of Information Systems* 16, 820 – 833.
- Cunningham, W. (2004). Wiki Design principles. Available on the web at: <http://c2.com/cgi/wiki?WikiDesignPrinciples> [Accessed September 2014].
- Dahlbolm, B. and Mathiassen, L. (1993). *Computers in context: The philosophy and practice of systems design*. NCC Blackwell, Oxford.
- Danis, C. and Singer, D. (2008). A wiki Instance in the enterprise: Opportunities, concerns, and reality. *Computer Supported Cooperative Work*, San Diego, USA.
- Eisenstadt, S. (1959) Bureaucracy, bureaucratization, and debureaucratization. *Administrative Science Quarterly* 4, 302 – 320.
- Grudin, J. and Poole, E. (2010). Wikis at work: Success factors and challenges for sustainability of enterprise wikis. In proceedings of the *6th International Symposium on Wikis and Open Collaboration*, Gdansk, Poland.
- Happel, H. and Treitz, M. (2008). Proliferation in enterprise wikis. In proceedings of the *8th International Conference on the Design of Cooperative Systems*, Carry le Rouet, France.
- Hasan, H. and Pfaff, C. (2007). Emergent conversational technologies that are democratizing information systems in organizations: The case of the corporate wiki. *Information Systems Foundations: theory, representation and reality*, 197 – 210.
- Hester, A. and Scott, J. (2008). A conceptual model of wiki diffusion. In proceedings of the *41st Hawaii International Conference on System Sciences*, Big Island, Hawaii, USA.

- Hildebrand C., Häubl G., Herrmann A., and Landwehr, J. (2013). When social media can be bad for you: Community feedback stifles consumer creativity and reduces satisfaction with self- designed products. *Information Systems Research* 24 (1), 14 – 29.
- Hodgson, L. (2004). Manufactured civil society: counting the cost. *Critical Social Policy*, 139 – 164.
- Holtzblatt, L., Damianos, L. and Weiss, D. (2010). Factors impeding wiki use in the enterprise: A case study. In Proceedings of the *Computer-Human Interaction*, ACM Publications, Atlanta, USA.
- Huang J., Baptista, J. and Galliers, R. (2013). Reconceptualizing Rhetorical Practices in Organizations: The Impact of Social Media on Internal Communications. *Information and Management*, (50), 112 – 124.
- Kallinikos, J. (2004). The social foundations of the bureaucratic order. *Organization* 11(1), 13 – 36.
- Kaplan A. and Haenlein, M. (2010). Users of the World Unite! The Challenges and Opportunities of Social Media. *Business Horizons*, (53), 59 – 68.
- Kärreman, D., and Alvesson, M. (2004). Cages in Tandem: Management, Control, Social Identity, and Identification in a Knowledge-Intensive Firm. *Organization* 11(1), 149 – 175.
- Klein, H. and Myers, M. (1999). A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. *MIS Quarterly*, 23 (1), 67 – 94.
- Kvale, S. (1996). *InterViews: an introduction to qualitative research interviewing*. Sage Publications.
- Majchrzak, A., Wagner, C. and Yates, D. (2006). Corporate wiki users: results of a survey. In Proceedings of WikiSym, Odense, Denmark.
- Majchrzak et al. (2013). The contradictory influence of social media affordances on online communal knowledge sharing. *Journal of Computer Mediated Communication*, 19 (1), 38 – 55.
- Mansour, O., Askenäs, L., and Abusalah, M. (2011). Wiki-based community collaboration in organizations. In proceedings of the 5th *International Conference on Communities and Technologies*, Brisbane, Australia.
- Mansour, O., Askenäs, L., Ghazawneh, A. (2013). Social Media and Organizing: An empirical analysis of the role of wiki affordances in organizing practices. In proceedings of the 34th *International Conference on Information Systems (ICIS2013)*, Milan, Italy.
- Martine, T., Cooren, F., Benel, A. and Zacklad, M. (2013). What does really matter in organizational wikis? A Sociomaterial approach. <sic 00821389>
- Miles, M. and Huberman, M. (1994). *Qualitative data analysis: an expanded source book*. Sage Publications, Thousands Oaks.
- Mintzberg, H. and McHugh, A. (1995). Strategy formation in an adhocracy. *Administrative Science Quarterly* (30), 160 – 197.
- Saldanha, T. and Krishnan, M. (2013). Organizational adoption of web 2.0 technologies: An Empirical analysis. *Journal of Organizational Computing and Electronic Commerce* 22 (4), 301 – 333.
- Schultze, U. and Avital, M. (2011). Designing interviews to generate rich Data for information Systems research. *Information and Organization* 21 (1), 1 – 16.
- Treem, J. and Leonardi, P. (2012). Social media use in organizations – Exploring the affordances of visibility, editability, persistence, and association. *Communication Yearbook* (36), 143 – 189.
- Vaast, E. (2007). What goes online comes offline: knowledge management system use in a soft bureaucracy. *Organization Studies* 28(3), 283 – 306.
- Wagner, C. (2006). Breaking the knowledge acquisition bottleneck through conversational knowledge management. *Information Resources Management Journal* 19 (1), 70 – 83.
- Weber M. (1978). *Economy and Society: An outline of interpretive sociology*. University of California Press, California.
- Yates, D., Wagner, C. and Majchrzak, A. (2010). Factors affecting shapers of wikis. *Journal of the American Society for Information Science and Technology* 61 (3), 543 – 554.

- Zammuto, R., Griffith, T., Majchrzak, A., Dougherty, D. and Faraj, S. (2007). Information Technology and the Changing Fabric of Organization. *Organization Science* 18 (59), 749 – 762.
- Zwass, V. (1992). *Management Information Systems*. Wm. C. Brown Publishers, Dubuque.

INTEGRATING OPEN EDUCATIONAL RESOURCES TO FOSTER SERIOUS GAMES AND GAMIFICATION DESIGN PRINCIPLES

Complete Research

*Kokkinaki, Angelika, University of Nicosia, Nicosia, Cyprus, kokkinaki.a@unic.ac.cy
Christoforos, Aresti, University of Nicosia, Nicosia, Cyprus, christoforos13@hotmail.com
Melanthiou, Yioula, University of Nicosia, Nicosia, Cyprus, melanthiou.y@unic.ac.cy*

Abstract

This paper presents aspects of the design, development and integration of open educational resources (OERs) on serious games design principles and gamification to foster good practices among faculty members of Higher Educational Institutions and other interested educators. The paper presents the methodology followed for the design and integration of OERs in this topic. The resulting modules are delivered through a synchronous e-learning platform and consist of thematic lectures, a variety of asynchronous tools for communication and support integration of OERs, serious games and gamification practices. Through its pilot implementation, learners became aware of aligning educational objectives to methodological principles and relevant tools using existing platforms, serious games and gamification aspects; they accessed videos and animations and they played games that let them realize the potentials of serious games in formal education or other learning contexts. Some preliminary results on their assessment are included.

Keywords: Open Educational Resources, Serious Games, Gamification, eLearning

1 Introduction

The research presented in this paper has been motivated by concerns of faculty members at Higher Educational Institutions (HEI) who had identified the need for effective means to promote students' motivation and engagement and eventually increase students' learning potential. As serious games and the selective use of gamification elements is a known approach to increase motivation, educators in all sorts of domains may wish to experiment with such approaches. However, not all of the educators are familiar with the required concepts and skills. Thus, the main objective of this research has been to use related theories and develop an e-learning platform that integrates OERs targeting educators who wish to employ serious games and gamification in learning activities designed for their students.

Integrating Open Educational resources (OERs) to support educators in their effort to undertake serious games and gamification elements for the development of learning activities is a complex task; it requires the alignment of learning objectives with learning tools and learning methods. In view of this, educators need to develop competencies on game elements and game mechanics, learn how to incorporate them within learning activities, which ought to be designed and developed with explicit consideration on the desired learning objectives for the end users.

Serious games, that is, applying games to non-game domains, have been extensively employed in educational contexts because they lead to high engagement of learners and through that to improvement of their skills. Recently, attention has been directed towards "gamification," which applies game elements and game design techniques to non-game systems (Deterding et al., 2011; Werbach and Hunter, 2012). As noted in (Cheong et al. 2013) a significant difference between the two concepts is that serious games are primarily games augmented with pedagogical components whereas

gamification refers to non-game activities to which game-like features have been added. In order to enhance the learning process and help learners conquer the specified learning outcomes, a gamified learning activity should be motivating, challenging, engaging and even stressful at times (Kapp 2012). This cannot be achieved by adding game elements like points, leaderboards, or badges to an existing activity or system; instead, it is the outcome of integrating game elements (Kapp, 2012; Nicholson, 2012) promoting a motivating and meaningful experience for the learners, especially suited for problem-solving and quick decision making. The volume and span of gamified applications (McGonigal, 2011; Reeves & Read, 2009) in business, education, health, marketing and other domains has also stirred interest among educators in the relevant fields and an inclination for some to acquire knowledge and skills that will enable them to employ it properly for pedagogical purposes in their courses. Despite the increased interest among educators, there has not been identified an e-learning initiative to examine how to educate course developers on this topic. To address this gap, an e-learning course for educators; it aims to provide them with the necessary knowledge and skills to identify learning objectives, design the learning activities and employ serious games and gamification to serve them. Based on the nationality of our audience the e-learning course content is in the Greek language.

Pedagogical, technological and economic aspects are involved in the design, development and implementation of e-learning courses and this renders the process challenging. For the specific course, considerations included the limited availability of resources in the specific language, the number of people who might be interested in and capable to follow this particular course, and the need for the methodology developed to be scalable.

This paper presents aspects of the design, development and implementation of an initial prototype of an e-learning course serious games and gamification, which was developed for Greek speaking educators who want to encompass serious games and gamification in their own courses. Through a localization process, however, other interested parties may localize the contents of the e-learning environment to their own special requirements. In an effort to preach what we pray, emphasis has been placed on employing game-based learning and appropriate gamification techniques for letting learners discover the different forms and aspects of serious games and gamification.

This paper aims at 1) outlining the methodology for identifying users educational needs related to the topic of gamification and serious games, 2) presenting a specific framework of the design of an e-course using game-based learning and gamification elements on the very same topic, 3) allowing interested parties who may wish to localize this course to understand the design parameters and thus perform their tasks more efficiently, and 4) facilitating the evaluation of the design and development of the e-course on serious games and gamification. For this course, the development team consisted of an expert on e-learning pedagogy and technologies with strong IS skills, a domain expert with experience and a developer. The development team was responsible for the design of the course and the corresponding learning material.

The remaining of this paper is structured as follows. Section 2 provides a brief literature review on relevant pedagogical approaches, game-based learning and gamification elements as well as the methodology used for identifying the educational needs of the target audience. Section 3 provides a description of the design of the e-course along with examples. Section 4 provides a roadmap through the implementation and evaluation of the environment on this implementation. Section 5 concludes this paper.

2 Learning Context and Identified Educational Needs

For the design framework of this research and the learning models to be applied, we follow a social-constructivist approach (Vygotsky, 1978), which proposes learner-centered activities and recognizes intrinsic learning through social interactions. Specific theories that guided the design of this e-course include:

- Vygotsky's (1978) theory of the "zone of proximal development", which posits that when learners engage in social behavior and enjoy the collaboration of peers they achieve more fulfilled learning experiences than what they might achieve by working individually;
- Kolb's (1984) experiential learning theory, which suggests that learning intertwined with applied, real life incidents and reflection upon this experience leads to higher learning achievements, and
- Lave's (1990) situated learning theory, which distinguishes a progression line among active learner initiating from "newcomer" status to "old timer" level within a learning community.

The overall design of the desired e-course observes is in alignment with these theoretical frameworks, furthermore, it employs game-based learning and other gamification elements.

A game may be defined as a system in which players engage in an abstract challenge, defined by rules, interactivity, and feedback, that results in a quantifiable outcome often eliciting an emotional reaction (Koster 2004). There is plenty of supporting evidence on the pedagogical role of fun in learning and of the potentials games to support learning (Ellis et al., 2006, Garris et al., 2002, Kolb 1984, Kolodner et al., 2003). Recently, gamification, that is, "[...] *the use of game design elements in non-game contexts*" (Deterding et al., 2011), has received attention as a potential mechanism for increasing users' engagement and behavioral change (Burke 2011).

The e-course was designed according to active learning, learner-centred principles and it includes certain common gamification elements and a set of learning activities that the learner performs in order to acquire or advance knowledge in this topic. The sequencing of these learning activities ranges from the simplest, to the most complex. The e-course specifies the learning objects involved (such as self-assessment test, video-lectures, exercises, handles to selected tools, text, games), the learning objectives (and other Metadata: who prepared it, when, for what type of learners, prerequisites, IPRs, technical infrastructure needed), and the learning environment (e.g., Moodle, WebEx, including the structure of the platform used, i.e. Moodle). Gamification elements are embedded in the e-course design; furthermore challenges, rewards, social influences and gamification specific self-expressing and self-assessment activities are included in the e-course.

In the design of the e-course we also included the profile of the learner, the learning objectives, possible prerequisites of each learning session, the learning environment structure and the utilized ICT skills. We have also included a description of the learning activities, possible sequencing of the learning activities and time restrictions or recommended timing: e.g., "one session per week". A (high level) description of the learning objects, including sizing (e.g., a 20min video), sources, and recourses necessary for implementation; as well as a series of self-assessment tools as well as guidelines for a cap-stone activity.

In alignment with the overall pedagogical paradigm and in order to identify the profile of the learners, the project team did secondary research. To maximize its potential impact on the target audience, it was designed as an open educational resource; the OER is in the Moodle server of the host University accessible at the The course includes discussions with facilitators, a variety of learning activities, methodologies and tools, relevant videos and animations and self-assessment exercises, as explained in the following section.

3 Design, Development and Integration of OERs

A fundamental principle followed in this research is the separation, with clearly defined interfaces, of the design, the implementation and the delivery of the course. The design requires the combined

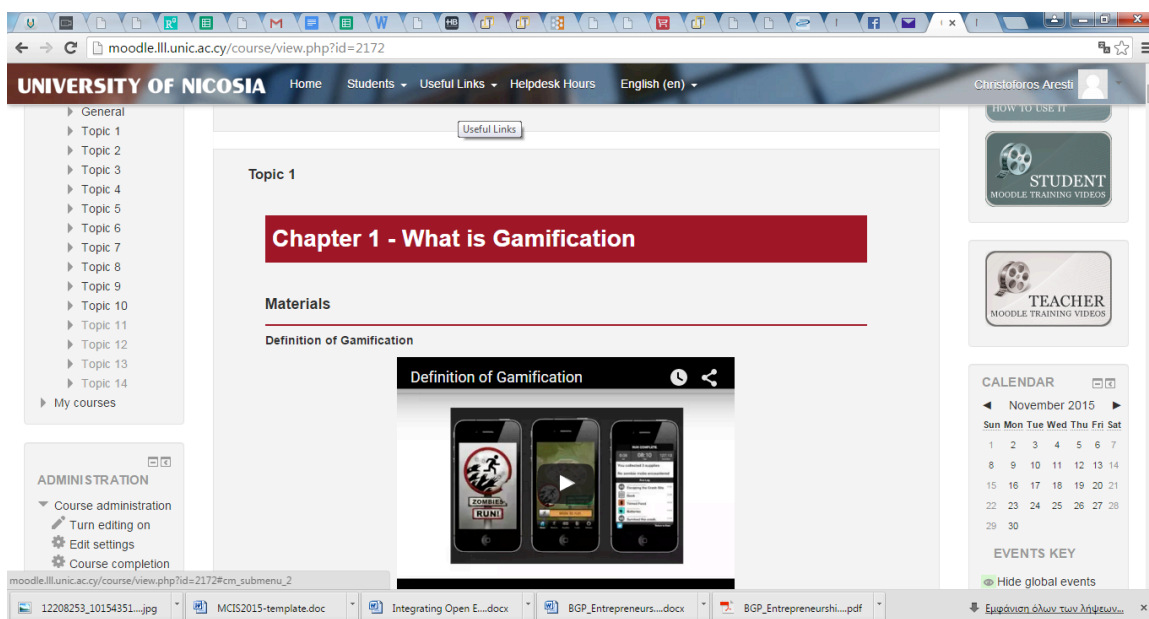


Figure 1. The Interface of the OER

expertise of e-learning pedagogy, technology and of domain knowledge. The implementation does not require the presence of an e-learning pedagogy or technology expert, any more. This separation of levels supports scalability and transferability.

The methodological approach allows the learners to perform a number of e-learning activities.

- Learners may participate in synchronous distance learning sessions that entail tutorial discussions on serious games and gamification concepts. Such sessions could be coordinated by a remote instructor. At implementation level, it is supported by WebEx®.
- An asynchronous e-learning system (Moodle, an open source LMS) and various Web and Web 2.0 tools were used. Furthermore, selected gamified activities have been developed and used as best practices.
- Participants design and develop their own educational material using appropriate method and tools (asynchronous distance learning).
- Learners may communicate, cooperate and form new knowledge with their fellow students, using embedded tools, educational activities and gamification elements.
- Finally, participants had the opportunity to undertake self-assessment exercises, play games and get engaged in gameful activities that empower learners follow their individual learning skills.

From a pedagogical perspective, the primary objectives of this course are to: i) raise awareness among learners on serious games and gamification; ii) let them develop skills and competences for developing their own educational resources undertaking appropriate design principles methods and tools and iii) enable them towards formative learning through individual projects. Example of how objectives (ii) and (iii) are implemented are given in the Figure 2 and 3.

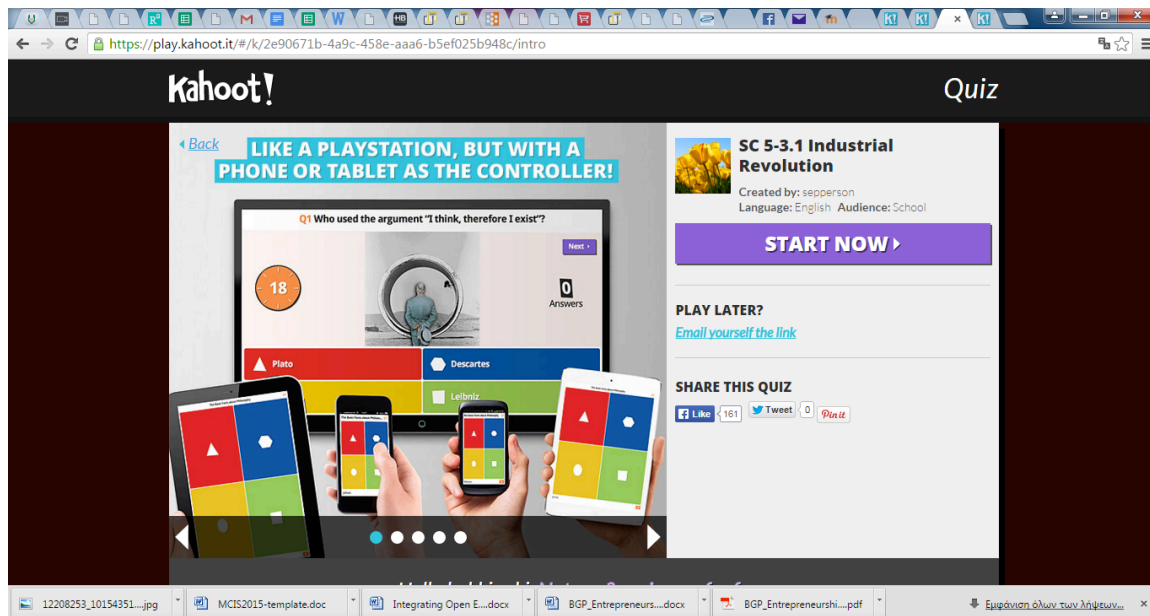


Figure 2. The educator may use an existing software tool for developing a serious game on the topic of his/her choice



[Επιστροφή στο Εκπαιδευτικό Περιβάλλον](#)

Figure 3. The educator may parameterize a well-known game and apply it in his/her course.

4 Assessment of the OER

We employed the Questionnaire for User Interaction Satisfaction (QUIS) for the assessment of the OER. Beyond the demographic questions, QUIS hierarchically organized measures of nine specific interface factors (screen factors, terminology and system feedback, learning factors, system capabilities, technical manuals, on-line tutorials, multimedia, teleconferencing, and software installation) as well as a measure of overall satisfaction. Each area measures the users' overall satisfaction with that facet of the interface, as well as the factors that make up that facet, on a 9-point scale. The questionnaire is designed to be configured according to the needs of each interface analysis by including only the sections that are of interest to the user. The QUIS questionnaire was translated into Greek and uploaded onto an online survey platform. The educators who had access to the OER were invited to access the system. The overall assessment of the project, including evaluation of the OER offered is described in more detail, as follows.

| 1. Γενική αντίδραση στο σύστημα | | | | | | | | | | | | |
|---------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-------------|-------------|-------------|-------------|-------------|---------------|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | N/A | |
| απεχθές | 2 2.3% | 5 5.7% | 4 4.6% | 2 2.3% | 1 1.1% | 3 3.4% | 9 10.3% | 9 10.3% | 16 18.4% | 27 31.0% | 9 10.3% | υπέροχο |
| δύσκολο | 2 2.3% | 3 3.5% | 1 1.2% | 0 0.0% | 4 4.7% | 1 1.2% | 11 12.8% | 7 8.1% | 17 19.8% | 30 34.9% | 10 11.6% | εύκολο |
| απογοητευτικό | 4 4.7% | 1 1.2% | 1 1.2% | 1 1.2% | 3 3.5% | 4 4.7% | 7 8.2% | 8 9.4% | 27 31.8% | 17 20.0% | 12 14.1% | ικανοποιητικό |
| ανεπαρκές | 1 1.2% | 3 3.7% | 1 1.2% | 3 3.7% | 4 4.9% | 2 2.4% | 7 8.5% | 10 12.2% | 29 35.4% | 14 17.1% | 8 9.8% | επαρκές |
| βαρετό | 3 3.7% | 1 1.2% | 1 1.2% | 2 2.4% | 2 2.4% | 6 7.3% | 8 9.8% | 8 9.8% | 20 24.4% | 16 19.5% | 15 18.3% | ενδιαφέρον |
| δύσκαμπτο | 3 3.7% | 3 3.7% | 0 0.0% | 2 2.5% | 2 2.5% | 5 6.2% | 13 16.0% | 22 27.2% | 16 19.8% | 11 13.6% | 4 4.9% | προσαρμόσιμο |

Figure 4. The overall assessment (part 1 of the QUIS questionnaire)

Overall, the e-course is very positively perceived by the learners; it also gathered high marks with regards to its perceived ease of use, satisfaction with regards to the scope of topics covered and completeness. The majority of learners found it interesting and flexible. Although the overall assessment is positive, the analysis of specific aspects reveals points for further improvement.

Follow up interviews validated that learners conquered the fundamentals on serious games and gamification, knew how to use tools to develop serious games and gamification, followed animation spots, conducted self-assessment tests, and overall enabled themselves to proceed in formative learning. Learners were particularly satisfied with regards to the methodology followed in the OER, although some concerns were expressed that the one-week time period between classes was not always enough for doing the suggested exercises and studying the additional material. Video clips and animations, were received very positively because they contained mostly applied examples, addressing all the main aspects of gamification. Another aspect of the integrated OERs that was very positively perceived was the gamified activities that were included.

In sum, the majority of the learners are satisfied with usability of the system, its content and its structure. Moreover, they are clearly in favor of the use of video segments and educational games. Users are content with the self-assessment exams, quizzes and the feedback given by fellow users and trainers. To a great extent, users believe that the OER provides the necessary content and skills to enable them undertake the serious games and gamification aspects in the design and development of their own educational material.

5 Conclusions

This paper focuses on the design and development of an OER on serious games and gamification. The authors report findings related to initial implementation; it is expected that the design process will go through repetitive iteration cycles to reach a higher level of maturity. The presented research works as a proof of concept that integrating OERs on serious games and gamification targeting educators can be an economically feasible and scalable process. The main contribution of such work remains to be validated through additional research. Another aspect that needs to be examined is the extent to which gamification has an impact on the learning process of users and explore ways that gamification can further support aspects of an overall educational strategy.

References

Burke, M. and Hiltbrand, T. (2011). How gamification will change business intelligence. *Business Intelligence Journal*, 16 (2), 8-16.

Ninth Mediterranean Conference on Information Systems, Samos, Greece 2015

- Cheong, C. Cheong, F. and Filippou, J. (2013). Using Design Science Research to Incorporate Gamification into Learning Activities. *PACIS 2013 Proceedings*. Paper 156. <http://aisel.aisnet.org/pacis2013/156>
- Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification". *Proceedings of the 15th International Academic MindTrek Conference*, 9–15.
- Ellis, J., Heppell, S., Kirriemuir, J., Krotoski, A., McFarlane, A. (2006). *Unlimited learning. Computer and Video Games in the Learning Landscape*, London: ELSPA Entertainment and Leisure Software Publishers Association.
- Garris, R., Ahlers, R. and Driskell, J. (2002). Games, motivation and learning: a research and practice model. *Simulation and Gaming*, 33, 441-467.
- Kapp, K. M. (2012). *The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education*. San Francisco: Pfeiffer.
- Kolb, D. A. (1984). *Experiential Learning*, Prentice-Hall, Inc., New Jersey.
- Koster, R. (2004). *A Theory of Fun for Game Design*, Paraglyph Press.
- McGonigal, J. (2011). *Reality Is Broken: Why Games Make Us Better and How They Can Change the World*. London: Penguin.
- Nicholson, S. (2012). *A User-Centered Theoretical Framework for Meaningful Gamification*. Paper presented at the Games+Learning+Society 8.0, Madison, WI, USA.
- Reeves, B., and Read, J.L., (2009). *Total Engagement: How Games and Virtual Worlds Are Changing the Way People Work and Businesses Compete*: Harvard Business Press.
- Vygotsky, L. S., (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press. Published originally in Russian in 1930.
- Werbach, K. and Hunter, D. (2012). *For the win*. Philadelphia: Wharton Digital Press.

KNOWLEDGE SHARING IN COMPUTERIZATION PROJECTS: AN APPROACH BASED ON BOUNDARY OBJECTS

Completed Research

Ben Chouikha, Mouna, InterVPNC Lab., University of Jendouba, Faculty of Law,
Economics, and Management, Jendouba, Tunisia, mouna.benchouikha@dauphine.fr

Dakhli, Salem Ben Dhaou, CERIA Lab., Paris-Dauphine University, France,
Salem.Dakhli@dauphine.fr

Abstract

Computerization projects play a critical role in providing applications that meet the information needs of organizations and facilitate organizational change. They function as a channel for the transmission of knowledge from individuals, methods, and past learning. As a result, knowledge management is a prerequisite to the success of such projects. Nevertheless, sharing knowledge held by organizational actors involved in computerization projects remains a difficult goal to achieve. In this paper, we analyze the obstacles to knowledge sharing in computerization projects in terms of knowledge boundaries. Apart from this analysis, the main contribution of this paper is the proposal of an approach to knowledge sharing in computerization projects, which relies on a typology of boundary objects. A case study based on development project of a reporting tool in an insurance company, has allowed us to verify the relevance of the approach proposed in this paper, and highlight its main managerial contributions..

Keywords: Computerization project, knowledge, knowledge boundary, boundary object, organizational actor, project space..

1 Introduction

The role of the information systems in the development and the survival of modern organizations is constantly growing (Henderson and Venkatraman, 1993; Henderson and Venkatraman, 1994; Hirschheim and Sabherwal, 2001; Sabherwal and Chan, 2001; Chan, 2001). Nowadays, information systems contribute to organizations efficiency and effectiveness and help them achieve difficult goals through providing an essential support to complex and interdependent organizational processes (Gurbaxani and Whang, 1991). The high frequency of changes induced by the external environment of organizations results in changes in the organizational processes and IT applications that support them. Not only such changes of information systems are difficult to implement but they are often related to severe cost, time and quality constraints. Moreover, the increasing share of the services economy (Horio and Watanabe, 2008) leads to institutional and legal changes that make value creation difficult without innovation of organizations catalogs of products and services and sales and delivery practices (Kotha, 1995; Kakati 2002). Innovation often results either in the creation of new organizational processes or in the modification of existing ones. To support efficiently innovative organizational processes, information systems must be agile and scalable. Moreover to be permanently effective and efficient, information systems evolves either by adding new applications or by modifying existing applications. Therefore, computerization projects play a critical role in providing applications that meet the information needs of organizations and facilitate organizational change. A computerization project is a temporary organization with unique and interdependent activities with a start date, an end date, a budget, and a goal consisting in the development of a software application to support partially or completely one or more organizational processes. The computerization projects can be considered as sets of organizational measures that enable modern organizations to reconcile short-term flexibility and long-term evolution, and constitute effective means to take into account new situations. These temporary organizations are complementary with permanent organizations that create stability (Ekstedt et al, 2002; Engwall, 2003; Manning, 2008; Kaulio, 2008). The success of computerization projects is a prerequisite to ensure that information systems play their full role in supporting organizational processes. It is measured by its efficiency and effectiveness. In other words, a computerization project can be considered a success if it results in an application that meets the needs of the organization, while respecting the constraints of cost, delay, and quality. However, successful computerization projects are rare as shown by many authors who noted that there is a software crisis born since the 60s of last century (Brooks, 1987, 1995; Toffolon, 1996; Toffolon, 1999; Reich, 2008). The software crisis has ramifications for both economic and social (Dakhli, 1998). Many solutions have been proposed in the literature to overcome the software crisis and reduce its impact. These solutions fall into two categories. The first category includes tools-based solutions aimed at improving the efficiency of computerization projects (Finlay and Mitchell, 1994; Coupe and Onodu, 1996). The second category comprises methods-based solutions intended to improving the effectiveness of computerization (Zachman, 1987; Sowa and Zachman, 1992; Boehm, 1988; Toffolon, 1996; Dakhli, 1998; Humphrey, 1999; Toffolon and Dakhli, 2002, Fayad et al., 2002; Long  p  , 2006; Ross, 2006; de Vaujany, 2009). While these solutions have made significant improvements both in terms of efficiency and effectiveness of computerization projects, they not eliminate all the symptoms and causes of the software crisis which persists and seems to be worsening. This failure is in our opinion due to the fact that the solutions proposed to date have ignored that knowledge is the essence of both the computerization process and artifacts produced by these processes. So far, computerization projects were analyzed as temporary organizations where the action is paramount and tasks, budgets, people, and schedules must be controlled to achieve the expected results. Despite its contributions, this perspective is gradually replaced by a complementary perspective which considers that computerization projects are temporary organizations where learning and knowledge are essential. According to this new perspective based on both theories of knowledge management and organizational learning, a computerization project functions as a channel for the transmission of knowledge from individuals, methods, and past learning (Walz et al., 1993; Faraj and Sproull, 2000;

Bredillet, 2004; Reich, 2007; Reich et al., 2008). During computerization projects life cycles, knowledge is shared, created, integrated, and operated to generate added value for organizations. Learning takes place for all stakeholders involved in a project of computerization: the sponsor, the project manager, business teams, and technical teams. Hence, the most important task of the project manager consists in combining knowledge from heterogeneous sources and mobilizing them to achieve the objectives of the computerization project he manages. As a result, management of the knowledge related to computerization projects is a prerequisite to deal with the software crisis. In particular, knowledge sharing by organizational actors involved in computerization projects is critical to the success of such projects. In this paper, we propose a knowledge sharing approach based on boundary objects. The goal of our work is to answer the following research questions:

- 1) What are the boundaries of knowledge in IT projects?
- 2) How can we manage the boundaries of knowledge in a computerization project?

To achieve this goal, we draw on the software global model proposed by Toffolon and Dakhli (2002) that we supplement by focusing on knowledge sharing. The motivation behind this choice of this model is related to its generic and structuring nature and its ability to describe the various aspects of computerization. Our paper is organized as follows. Section 1 explains why knowledge is essential for computerization projects. In Section 2, we present the theoretical foundations of our work comprising on the one hand, a synthetic description of the software global model and on the other hand, a presentation of the boundary object concept. In Section 3, we analyze how the management of knowledge boundaries contributes to knowledge sharing. Section 4 uses the software global model to identify the boundaries of knowledge in computerization projects. The boundary objects to manage these boundaries are described in Section 5. In Section 6, we present our research methodology, which includes a case study described and analyzed using the approach proposed in this paper. Section 7 concludes the paper by listing the contributions and limitations of the proposed approach, and future research directions.

2 Computerization projects knowledge

The computerization projects are based on knowledge-intensive activities and result in informational artifacts (applications, documents, models, ...) that can be considered as an accumulation of knowledge shared by the organizational actors involved in these projects (Baetjer, 1998). According to this author, the effectiveness of these artifacts increases with the amount of shared knowledge between their stakeholders. Koskinen et al. (2003) and Newell et al. (2004) have confirmed that knowledge is created and shared in computerization projects. Other authors consider that computerization projects are temporary organizations that rely heavily on knowledge held by their team members to create informational artifacts (Lundin and Soderholm, 1995; Packendorff, 1995; Sodedund, 2002; Turner and Muller, 2003). Computerization projects team members learn both individually and collectively, transfer their knowledge to others, and create new shared concepts. Thus, computerization projects management and knowledge management are two nested processes (Leseure and Brookes, 2004). Moreover, Becker (2001) noted that the success of computerization project requires integrating knowledge distributed among several individuals and groups within the organization. Other authors consider that social capital is a possible channel for integration of organizational knowledge dispersed across individuals (Newell et al, 2004. Chiu et al, 2006). According to Adler and Kwon (2002), social capital is the capital sympathy resulting from the construction of social relations that can be mobilized for supporting action. Putnam (2000) identified two types of social capital: bonding social capital and bridging social capital. The former is a glue that strengthens the cohesion of teams through internal relationships between their members while the later relies on relationships between organizational actors to enhance the cohesion of social groups who do not know each other before. Newell et al. (2004) analyzed the impact of social capital on the integration of knowledge throughout a computerization project. The results of their work distinguish the initial phase of the other phases of a

computerization project. During the initial phase of a computerization project, the members of this project team use bonding social capital to weld the team and strengthen its internal cohesion by creating a common understanding of the project objectives and sharing their knowledge. As the project progresses, the members of the project team mobilize bridging social capital to acquire external knowledge distributed among several social groups in order to achieve the project objectives. The results of conceptual work emphasizing the importance of learning and knowledge in computerization projects have been confirmed by several empirical studies. For example, Faraj and Sproull (2000) pointed out that the coordination of experts facilitates project members learning and improves their performance. Yoo and Kanawattanachai (2001) concluded that the success of a computerization project depends firstly on the knowledge held by each member of the project team about the expertise areas of other team members and secondly on the ability of each member of the project team to mobilize knowledge to achieve the project objectives. Knowledge relevant to computerization projects fall into four categories identified in the literature related to organizational learning and knowledge management (Reich and Wee, 2006; Reich, 2007): process knowledge, domain knowledge, institutional knowledge, and cultural knowledge. The process knowledge refers to knowledge held by computerization project stakeholders about the project structure, tasks, milestones, and methodology (Chan and Rosemann, 2001; Meehan and Richardson, 2002). It is a body of knowledge that makes computerization project stakeholders aware of their contribution to the project and what is expected of them in terms of deliverables, and helps them organize themselves to achieve the project's goals. Domain knowledge includes business knowledge, organization's context knowledge (opportunities, problems, potential solutions, ...), and products related knowledge (Chan and Rosemann, 2001). The domain knowledge is dispersed throughout the organization. For example, the sponsor of a computerization project often holds the most important knowledge about the business, the problem to solve and its potential solutions, and the existing opportunities. Technical experts within and outside the organization have the knowledge related to the most appropriate technologies to support the possible solutions to the computerization problem. Likewise, members of the project team have a deep understanding of the organization and its organizational processes impacted by the computerization project. Institutional knowledge are a mixture of stories, information about power structures, and values of the organization. They don't relate to facts but focus on how to interpret and understand facts. This category of knowledge is essential especially for a technology provider or an external project manager to understand the computerization and business problems, their solutions, and the decisions taken as part of the computerization project. Cultural knowledge is based on both the areas of expertise and the national cultures of the computerization project team members.

3 Theoretical foundations

In this section, we present synthetically the boundary object concept and the software global model which constitute the theoretical foundation of our work. In particular, we use the software global model to identify the computerization projects knowledge boundaries.

3.1 The software global model

Based on the economic agency and transaction costs theories (Coase, 1937; Williamson, 1981; Alchian and Demsetz 1972; Jensen and Meckling, 1976) and the theory of dimensions of the software (Toffolon, 1999), the software global model is a framework for software production in modern organizations (Toffolon and Dakhli, 2002). According to this model, a computerization project is a temporary organization characterised by a structure where a group of individuals perform tasks in accordance with a production technology to achieve a set of objectives (Leavitt, 1963). This model considers that the structure of a computerization project is based on four project spaces: the problem space, the solution space, the construction space, and the operation space. The computerization project stakeholders are represented by four types of organizational actors: the customer, the architect, the developer, and the end user. Each type of actor is associated with a project space where he plays the

principal role and may play a secondary role in other project spaces. The customer is associated with the problem space and plays a secondary role in the solution and the operating spaces. The architect is associated with the solution space and plays a secondary role in the problem and the construction spaces. The developer is associated with the construction space and plays a secondary role in the solution and the operating spaces. The end user is associated with the operation space and plays a secondary role in the construction and the problem spaces. The software global model describes computerization projects as network of producer-consumer contracts between organizational actors belonging to the four types listed above. These contracts are realized according to a meta-iterative process called PACO (Problem - Solution - Construction - Operation) which characterizes the dynamic part of a computerization project i.e. the tasks performed by the members of the project team. A computerization project starts from an organizational problem identified and resolved in the problem space. The organizational solution is submitted by the customer to the architect who defines, in the solution space, a software solution (software architecture) to computerize the organizational solution. This solution is implemented in the construction space by the developer who delivers a software which supports the activities carried out by the end users in the operation space. In this paper, we complete the dynamic part of the software global model by adding a project management process implemented by a project manager who plays the principal role in a project management space. The project manager plays a secondary role in the four project spaces which characterize the static part of the original software global model. Likewise, the four types of organizational actors involved in a computerization project play a secondary role in the project management space. The project manager is the contact for all other organizational actors involved in the computerization project, and the guarantor of the projects efficiency and effectiveness. Finally, in each of the five project spaces, the organizational actors involved in a computerization project use a production technology to accomplish their tasks and achieve the project goals. This technology consists of material resources, methods, techniques, tools, and existing procedures within the organization. The following diagram (Figure 1) illustrates the computerization project model used in this paper.

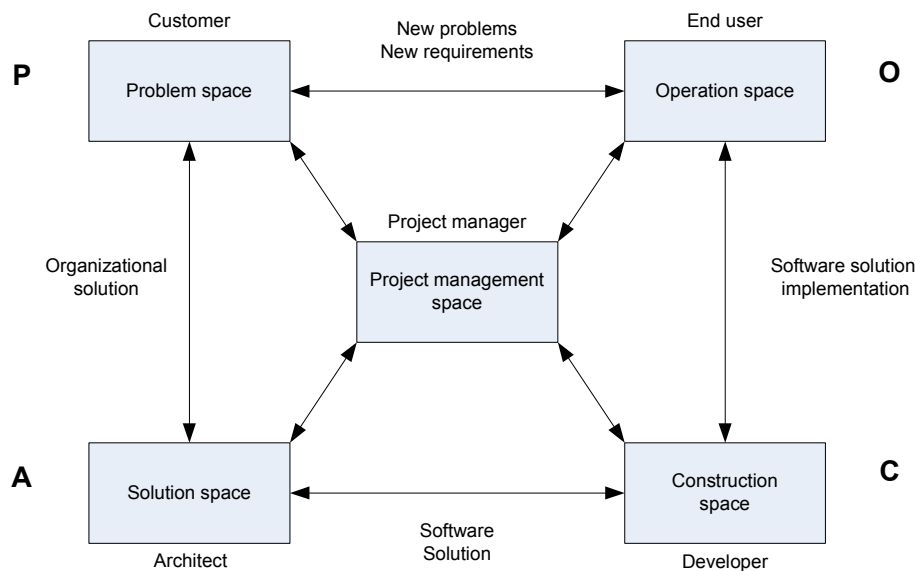


Figure 1. The computerization project model

3.2 The boundary object concept

The concept of boundary object was introduced by Star and Griesemer (1989) as an instrument of cohesion of many social worlds around a single shared goal. According to these authors, a boundary object is positioned at the crossroads of several social worlds to take into account the needs and

constraints of each world. It is an instrument with a common set of shared knowledge that facilitates understanding and cooperation, and minimizes conflicts between organizational actors who have different representations of a problem. Thus, the boundary objects are characterized by their ability to federate, around a common goal, organizational actors with heterogeneous identities and different or conflicting interests. In other words, boundary objects are used for information collection, and knowledge sharing between organizational actors belonging to different social groups. Whether general or specific to a social world, a boundary object has three essential characteristics identifiable by all the social groups using it. First, it is flexible enough to adapt to the needs and constraints of organizational actors from different social worlds. Secondly, it is robust enough to maintain a common identity among many social worlds. Finally, a boundary object is not very structured in order to make its use easier in different social worlds. In addition to the three essential characteristics listed above, Vinck (2009) considers that a boundary object must be modular, multipurpose, standard, and enough abstract. In particular, the multipurpose nature of a boundary object means it can be used in the context of many different activities and practices. Star and Griesemer (1989) distinguish four categories of boundary objects: repositories, ideal types, multi-perspective objects, and standard forms. Repositories contain data that can be used by different organizational actors based on the social world they come from and their own practices. Databases and intranets are examples of repositories. Ideal types are boundary objects whose content is not identical for all social groups who adapt information inherent in these objects according to their needs and their particular contexts. Ideal types include notably computerization methods which are customized by organizations applying them. Multi-perspective boundary objects are shared by different social groups but their contents and their roles are perceived differently by these groups. For example, a software architecture model is a multi-perspective boundary object which is considered as a constraint and a source of delay by the project managers, and as a prerequisite for the quality and sustainability of the software solution provided by the information system architects. The standard forms are tools for the collection, aggregation, and analysis of data in a format shared by different organizational actors. The dashboard of a computerization project is an example of standard form shared by all project stakeholders. Briers and Chua (2001) suggest a fifth boundary objects category characterized by a high level of legitimacy and shared by a significant number of organizational actors either intra or inter-organizations that can use them specifically. Business rules, best practices shared by many managers, and organizational culture fall into this category. Organizational culture is a category of boundary object can be generalized to include moral and cultural values shared by organizational actors belonging to different social groups.

3.3 The management of knowledge boundaries'

The work of Carlile (2002, 2004) about the role of boundary objects in knowledge sharing is among the main contributions in this field. The significance of boundaries between different social groups and different disciplines has been highlighted by several authors (Leonard-Barton, 1995; Brown and Duguid, 2001). Based on the findings of Star and Griesemer (1989), Carlile (2002) identified three relational properties of knowledge at the boundaries between different fields. These properties are the difference, the dependence, and the novelty (Rebentisch and Carlile, 2003). The difference results from of the specialization areas and for the amount or type of accumulated knowledge. The concept of dependence between two entities defined by Litwak and Hylton (1962) as the condition requiring the two entities to consider each other to achieve their goals. The third relational property of knowledge at boundaries is related to the degree of novelty which impact the relations between actors belonging to different social worlds and involved in a project. Carlile (2002) noted that the emergence of a novelty at the knowledge boundaries between different disciplines or different social groups can result in the lack of a common language for representing the differences and dependencies between actors involved in a project.

According Carlile (2002, 2004, 2005), the management of knowledge boundaries is necessary to facilitate knowledge sharing in order to achieve the common goals of project team. This author

distinguishes three types of knowledge boundaries (syntactic, semantic, and pragmatic) and proposes three approaches to manage efficiently Carlile (2004, 2005).

The syntactic boundary is characterized by the stability and the control of the differences and the dependencies between the organizational actors. Carlile (2004, 2005) stresses that knowledge transfer is the most appropriate approach for manager syntactic boundaries. Such an approach - called oriented information processing - is based on the development of a common language to share and evaluate knowledge at the boundaries.

The transition between the syntactic and the semantic boundaries occurs when a novelty makes ambiguous at least partly knowledge dependencies, differences or interpretations. To manage the knowledge semantic boundaries, Carlile (2004) suggests a translation approach called interpretative approach for creating common interpretations that constitute adequate means to share and evaluate knowledge at the boundaries.

The transition from a semantic pragmatic boundary to a pragmatic boundary occurs when an important novelty generates conflicting interests between actors. These interests constitute barriers to sharing and assessment of knowledge across boundaries. Indeed, in case of conflicting interests, knowledge developed in one domain often generates negative consequences for other domains, materialized by costs incurred for the actors belonging to these domains. Because of these costs, the actors are less willing to make the changes induced by an important novelty. According Carlile (2004), the management pragmatic boundaries is based on a transformation approach - called political approach - which facilitates the development of common interests through the transformation of actors conflicting interests and knowledge. These common interests constitute adequate means to share and evaluate knowledge at boundaries.

Carlile (2004) suggests that knowledge sharing at boundaries requires the creation of boundary objects common to organizational actors involved in a project. The knowledge representation using boundary objects (Star 1989) and the knowledge management process at boundaries that these objects facilitate is the key to the development of an effective context common to organizational actors who hold different types of knowledge. According to Carlile (2002), effective boundary objects establish a common language for representing knowledge and provide a concrete method to specify differences and dependencies of knowledge at boundaries. Boundary objects can be technical specialists (individuals), mutually accepted methods, or shared artifacts. Whatever their category, they not only facilitate the representation of knowledge, but also its translation and transformation. Moreover, the boundary objects provide guidance regarding the roles of the actors involved (who defines the knowledge representation? Who can modify it?). Ben Chouikha (2010) confirmed Carlile's conclusions (2002, 2004) while analyzing the design of earning organizations and validated these conclusions using a longitudinal case study focusing on several project teams in a context merger between two international companies of telecommunications services. Ben Chouikha and Dakhli (2011) completed Carlile's work through a study knowledge sharing in a virtual organization.

The creation of boundary objects for management of knowledge at boundaries leads to transformation of knowledge held by different actors interacting at boundaries. Such transformation starts when relevant sources of knowledge have been identified. If the degree of novelty is low or the dependencies between specialization domains are stable with defined boundaries or interfaces, the transfer of knowledge may be sufficient for knowledge sharing between individuals, groups, or organizations (Argote, 1999; Winter and Szulanski, 2001). However, the innovation degree increases, the differences and dependencies between individuals and groups of individuals often generate negative effects and problems that must be solved jointly. Thus, the specialized knowledge and expertise to be integrated must be transformed in order to eliminate the negative effects, and build a collective solution. The challenge of creating new knowledge is not limited to the transition from tacit knowledge to explicit knowledge through externalization as noted by Nonaka and Takeuchi (1995) but consists in negotiating, redefining, transforming knowledge used by different groups to create a new solution (Carlile, 2002).

We note that the representation of knowledge, which plays an important role in this step, takes place at two levels: intra-specialization and inter-specializations. In particular, the bottleneck of the transformation of knowledge is through the different specialization domains (inter-specializations) since experts in different domains lack a common language or - in the case of complex situations - a common method of negotiation and trade-off between various alternative possibilities that enable different sources of knowledge to contribute to the construction of a new product or service. In other words, experts who hold different knowledge types need a shared context offering tangible and relevant common method and means to specify their differences and their dependencies. In particular, since the knowledge transformation process is iterative, this common method must improve the experts ability to represent, specify and negotiate their knowledge transformation throughout the iterations (Carlile, 2003). As a result, the integration of knowledge held by several organizational actors presents many difficulties and depends on the context in which these actors operate (Becker, 2001; Maaninen-Olsson et al., 2008; Carton and Farastier, 2012).

4 The knowledge boundaries in computerization projects

In this section, we first show that the difficulty to share knowledge in computerization projects is due to the existence of many knowledge boundaries between actors involved in such projects. Then we use the software global model to develop a typology of knowledge boundaries in computerization projects.

4.1 Knowledge sharing barriers in computerized projects

Becker (2001) has highlighted the difficulty to manage knowledge distributed among several individuals and groups within an organization. In particular, the computerization projects involve actors whose roles are multiple and involvement is variable. The interactions between these actors - needed to achieve the project objectives - are impacted by the specific nature of the organizational context of the project (cost, quality, time, ...). The dispersion of organizational knowledge may result in inadequate mobilization of critical knowledge due to differences in status between the project team members, the physical distance separating them, or the lack of familiarity related to differences in culture, thinking patterns, and languages. This is confirmed by Reich (2007) who has established that there is a lack of common understanding of knowledge management in the context of computerization projects and that the project team knowledge is difficult to mobilize to achieve the project objectives because of the absence of a knowledge management approach implemented within the organization (Arthur et al., 2001). According to Reich (2007), the knowledge management in the context of a computerization project consist in applying a set of principles and processes designed to bring significant knowledge to the project team. Thus, the effective management of knowledge facilitates the creation and integration of knowledge, minimizes the amount of lost knowledge, and reduces knowledge gaps between the members of the project team.

Hardy et al. (2005) defined the effective cooperation as a process based on the differences between the participating actors to produce innovative solutions that represent a consensus among the concerns and viewpoints of these actors. This process is facilitated by the existence of an identity and shared practices (Hardy et al., 2005; Levina and Vaast, 2005; Feng et al., 2011). However, the differences in status between actors involved in cooperation constitute barriers for this process (Hoegl and Gemuenden, 2001; Levina, 2005; Metiu, 2006).

The ineffectiveness of cooperation in computerization projects is due to the existence of social or organizational boundaries between the members of the project team, which impede the creation of a common identity and shared practices and reinforce the status differences. Spatial or temporal physical distance between the members of project team is an example of organizational boundary (Levina and Vaast, 2005, 2008). Similarly, differences between spoken and written language, national or professional cultures, and the positions held are social boundaries that can both exacerbate inequality and reinforce status differences between the members of project team (Lam 1997; Espinosa et al.,

2003; Krishna et al 2004;. Levina and Vaast, 2005; Levina and Vaast, 2006; Birnholtz and Finholt, 2007; Cramton and Hinds, 2007; Levina and Vaast, 2008).

Apart from organizational and social boundaries, Cramton and Hinds (2007) and Walsham (2002) emphasized that the boundaries resulting from the differences in practices between the members of project team are the most important barriers that inhibit the cooperation between these actors. Cooperation in a computerization project is made even more difficult by the fact that a significant part of the knowledge necessary to the project is tacit or embedded into the practices of the different actors involved in this project (Lam, 1997; Cramton, 2001; Hinds and Mortensen, 2004; Metiu, 2006, Levina and Vaast, 2008). Metiu (2006) pointed out that the status differences between computerization project team members can emerge from the context of project: conflicts related to design patterns or computer programs property,... Furthermore, the differences between the professional and industrial practices result in more knowledge boundaries that can reinforce status differences by creating new power relationships, and prevent the integration of new members or external stakeholders in a computerization project team (Bourdieu, 1984; Carlile, 2004; DiMaggio, 1991; Montgomery and Oliver, 2007).

4.2 Typology of knowledge boundaries in computerization projects

The software global model (Toffolon and Dakhli, 2002) shows that computerization projects take place in five project spaces (four areas of production and project management space) in which five types of organizational actors play either a key role either a secondary role. Software production is consistent with the iterative metamodel designated by the acronym PACO (Figure 1). Thus, a computerization project consists of three parts: static, dynamic, and organizational. This results in three types of knowledge boundaries:

- inter-spaces production knowledge boundaries,
- intra-spaces production knowledge boundaries,
- knowledge boundaries between between the project management space ant the organization including the production spaces.

The production inter-spaces knowledge boundaries result from the fact that the four types of project actors associated with the four spaces have specific and specialized knowledge. Thus, customer knowledge is related to his profession and to the problem and organizational solution to be computerizes. We note that even if the customer has knowledge about information and communication technologies, this knowledge is often too general. Similarly, the architect has only shallow and general knowledge of other businesses including those of the customer and the end user. Many architects know only little about development languages and databases. This is the case, for example, of functional architects. Moreover, developers usually ignore architectural rules and often consider them as constraints. They also generally have difficulty understanding the requests and questions of end-users who try to communicate with them without using a computer language.

The production intra-spaces knowledge boundaries result from the fact that a type of organizational actors associated with a production space includes a heterogeneous population of organizational actors holding different and specialized knowledge. For example, the "architect" type includes all organizational actors providing transverse support for computerization projects like functional architects, technical architects (software architecture specialists, infrastructure architects), experts in methodology, quality engineers, and databases experts. This is also the case of "developer" type which includes analysts, designers, programmers, and software integrators. Similarly, the "end user" type can include organizational actors belonging to different hierarchical levels and specialized in different businesses Finally, the "customer" type refers to a heterogeneous population consisting of middle managers, senior managers, and executives of an organization.

The knowledge boundaries between the project management space and the rest of the organization are related to the needs of communication and coordination of the project manager not only with

organizational actors from the four production spaces, but also with the computerization project stakeholders. First, the project manager must interact with the customers to understand their needs and to report on the progress of the project. Second, he must also interact with the architect who helps him define a sustainable IT solution supporting the organizational solution. Third, the project manager must monitor the progress of their project, coordinate the project team and manage priorities, conflicts and resources. Finally, he must report to the project governance authorities. Thus, the project manager is in constant interaction with a heterogeneous population of organizational actors he does not share the specialized knowledge and expertise. It follows that computerization projects knowledge boundaries may be syntactic, semantic or pragmatic. In the remainder of this paper, the knowledge boundaries between the project management space and the organization are designated by the term "project management boundaries". The following table (Table 1) summarizes the computerization project types of knowledge boundaries.

| Type | Boundaries list |
|---------------------------|--|
| Inter-spaces | - Customer - Architect - Architect - Developer - Developer – End user - End user - Customer |
| Intra-space | - Architect - Architect - Developer - Developer - End user – End user - Customer - Customer |
| Project management | - Project manager - Customer - Project manager – Architect - Project manager – Developer - Project manager – End user - Project manager – Governance authorities |

Table 1. *Typology of knowledge boundaries in computerization projects*

5 The management of computerization projects knowledge boundaries

To manage boundaries knowledge in computerization projects, various boundary objects can be defined. Prior to listing examples of boundary objects, we propose a classification of these objects according to their main characteristics.

First, the three properties of boundary knowledge - the difference, the dependence, and the novelty - identified by Carlile (2002) impact the producer/consumer contracts characterizing the relationships between organizational actors involved in computer projects. As a result, the boundary objects facilitating boundary knowledge management play an important role in improving the development and implementation of these contracts. Thus, a boundary object can have four roles: transfer, clarification, production, and communication.

In a computerization project, a transfer boundary object provides the actors involved in a syntactic knowledge boundary with a common language which helps them transfer the knowledge necessary to carry out the contract between them. A set of specifications for the development of accounting software is an example of transfer boundary object. A clarification boundary object reduces uncertainty and ambiguity inherent in a contract between organizational actors involved in a semantic boundary. This is the case, for example, of an informative prototype produced by the requirements engineering process (Toffolon and Dakhli, 2002). We note that a clarification boundary object may be an organizational actor. For example, a facilitator is a clarification boundary object as it helps the architect and the developer reduce the ambiguity inherent in the organizational solution to computerize

Similarly, an architect immersed in a computerization project is a clarification boundary object as it helps developers understand the solution to implement, and the rules and architectural constraints to be taken into account. A production boundary object is a final artifact accepted by organizational actors participating in a knowledge boundary and preserved during the software lifecycle. For example, an architecture report validated by the customer and accepted by the project manager and the developer may be considered as a production boundary object. This is also the case of a software version accepted by the customer and the end-user who implements it in the operating space. Generally, production boundary objects permit managing pragmatic knowledge boundaries. However, such boundary objects can be used to manage a syntactic boundary in particular in the case of lack of novelty.

A communication boundary object enables managing project management boundaries which involve the project manager. Indeed, the project manager must communicate with various organizational actors concerned with the computerization project. Considering the heterogeneity of these actors and knowledge they hold, the project manager uses various communication boundary objects for management of project boundaries. For example, a simplified Gantt chart is a suitable communication boundary object for management of the Project manager – Customer boundary while a detailed GANTT chart, or a PERT chart are communication boundary objects for management of the Project manager – Developer boundary. Similarly, a dashboard describing synthetically the progress of the project and explaining the delays is a communication boundary object which help manage the Project manager border - governance authorities boundary. We note that the communication boundary objects are generally standardized and help manage the three kinds of knowledge boundaries.

Second, the boundary objects are derived from the three approaches proposed by Carlile (2004) depending on the knowledge boundaries to be managed (Feng et al., 2011). Thus, it is obvious that the transfer boundary objects are derived from the knowledge transfer approach. Similarly, the clarification boundary objects result from the interpretive approach. Moreover, the production boundary objects are often issued from the political approach. Indeed, such boundary objects are finished products which must reconcile the conflicting interests of organizational actors involved in a pragmatic boundary. However, if there is no novelty, a production boundary object can be derived from a knowledge transfer approach. Communication boundary objects may be derived from one of from the three approaches proposed by Carlile (2004). Indeed, since the content of communication boundary objects often includes indicators or instruments related to intangible artifacts, these objects should be consensual to reflect the conflicting viewpoints of the computerization project stakeholders. In this case, they may be considered the result of the political approach. In other cases, the communication boundary objects are derived from the interpretive approach. This is the case, for example, if an organizational actors involved in a knowledge frontier does not understand the meaning of the content of a communication boundary object. However, many standards of communication boundary objects exist and are accepted by most of the organizational actors involved in computerization projects. The relationships between knowledge boundaries management approaches (Carlile, 2004) and boundary objects are summarized in Table 2.

Finally, a boundary object can be described using three dimensions: boundary, role, and knowledge boundaries management approach. This provides a typology consisting of a large number of boundary objects associated with a computerization project. Certainly, all these boundary objects do not have the same weight in the management of knowledge boundaries of a computerization project. Nevertheless, that their number is important is an indicator of the complexity of knowledge boundaries and the difficulty of sharing knowledge in a computerization project team. Table 3 provides examples of boundary objects used in computerization projects.

| Boundary object | Management approach of knowledge boundaries |
|--------------------------------------|---|
| Transfert boundary object | - Knowledge transfert approach |
| Clarification boundary object | - Interpretive approach |
| Production boundary object | - Political approach - Knowledge transfert approach |
| Communication boundary object | - Political approach - Interpretive approach - Knowledge transfert approach |

Table 2: Relationships between knowledge boundaries management approaches (Carlile, 2004) and boundary objects

| Boundary | Example of boundary object | Role |
|---|---|---|
| Customer - Architect | - Specifications - Facilitator - Architecture report | - Transfert - Clarification - Production |
| Architect - Developer | - List of architecture rules to be applied - Architect immersed in the project team - Access to architecture FAQ - Architecture report | - Transfert - Clarification - Clarification - Production |
| Developer – End user | - Informative prototype - Mock up - Training session - Final software version | - Clarification - Clarification - Clarification - Production |
| End user - Customer | - Requirements list - Software evolution requests list | - Transfert - Transfert |
| Project manager – Governance authorities | - Project status dashboard | - Communication |
| Project manager - Developer | - Detailed GANTT chart - Reference schedule - Meeting report | - Communication - Communication - Communication |
| Developer - Developer | - Data model - Tested program to be integrated - Access to development FAQ - Access to a repository of coding problems and their solutions | - Production - Production - Clarification - Clarification |

Table 3: Examples of boundary objects used in computerization projects

6 Research methodology

In this section, we briefly present the case study, the data collection methodology, and the analysis of results.

6.1 The case study description

The empirical validation of the approach presented in this article is based on a case study based on a development project of a reporting tool in an international insurance group. This group was formed through many mergers and acquisitions involving French and foreign insurance companies. It now brings together about 30,000 employees in several countries. The strategy of this group is to significantly increase its share in the savings market. The target customer segment consists of affluent

customers able to save between 100,000 € and 500,000 € excluding real estate. To capture these customers, the insurance group sets up a specific commercial approach and dedicated sales forces. In order to enable financial advisors to offer these customers the best quality of service, it is planned to provide them with a comprehensive reporting tool for monitoring saving and customers assets. The purpose of this tool - designated in the remainder of this paper by the acronym NREP (New Reporting) - is to provide financial advisors with a consolidated view of the assets of each customer, facilitate financial analysis of the customer assets portfolio, and provide each customer with a reporting on the status of his account. Similarly, the insurance group's marketing department has expressed interest in NREP tool and expect that this tool provides indicators that facilitate the implementation of marketing campaigns. In other words, the main functions of the reporting NREP tool consist in:

- providing customers with a high quality after-sales service,
- enabling the financial advisors and the customers share the same objective vision of saving contracts,
- enabling a delegation – to each the customer - of some acts of management related to its contracts (payment, buyout,...),
- help financial advisors preparing their interviews with customers,
- reduce the customer's telephone calls to his financial advisor,
- facilitate the framing and focus of marketing campaigns.

The NREP tool will be accessible to the customers so that their contracts are visible at any time. The process to be computerized describes the commercial approach followed by financial advisors to capture affluent customers targeted by the insurance group strategy. This process has four steps. The first step consists in making a diagnosis of the customer's wealth. This diagnosis is the basis from which the financial advisor defines the advice that he will propose to the customer. The second step perform simulations are performed to convince the customer by demonstrating the relevance and appropriateness of this advice. The third step consists in developing a reporting accessible at any time by the customer and the financial advisor in order to ensure the transparency of the relationships between the advisor and the customer. The fourth step is carried out by financial advisors and consists in monitoring the customers assets portfolio. Currently, financial advisors of the insurance group do not have a single, comprehensive view of all savings contracts of their customers. Certainly, many reporting tools are available within this insurance group and used by financial advisors. However, these tools do not allow financial advisors consolidate the results of their customers portfolios, present the following several analysis axes, and querying the customers portfolios and contracts using several criteria.

The tool NREP to be developed is a data container for financial advisors, customers and savings contracts. Data is managed by existing IT applications belonging to the insurance group's information system.

After a preliminary study, the insurance group to develop the NREP tool in house by its French subsidiary Information Systems Departement composed of an Architecture unit and a Development unit. The Architecture unit consists of functional architects, technical architects, methods experts, quality engineers, databases engineers, and experts in reusable technical components. These human resources are either employees of the insurance group or external consultants. The Development unit consists of analysts and designers who are either employees of the group or external consultants. Implementation, unit testing and integration testing of applications are carried out by an Indian "off-shore" company. Software applications developed in house must respect the Information System architecture and urbanization rules published in a guide available to the Development unit. The computerization projects should be conducted according to a development method defined by the methods engineers. Despite this method is based on the waterfall lifecycle (Boehm, 1976, 1984), it enables software prototyping during the user's requirements definition phase and allows iterative

development according to the spiral model (Boehm, 1988). Apart from the geographical and cultural dispersion of the NREP computerization project design and implementation teams, this project has two important challenges. On the one hand, users have not formally expressed their needs and requirements. On the other hand, to take into account the diversity of savings contracts proposed by the insurance group, the NREP tool must be able to exchange informational flows with all applications that manage these contracts. It follows that the NREP tool must take into account the technical constraints (database structure, data format, interface constraints, ...) and the operational constraints (information availability, frequency of exchange, ...) of these applications.

6.2 Data collection and results analysis

The data collected is issued from two sources among those identified by Yin (1994). On the one hand, we have consulted the documents made available by the project manager or published on the project intranet, the documents available on the intranet site of the French subsidiary of the insurance group or in the documentation repository, and the documents provided by the Architecture unit. On the other hand, we have collected information on the project progress and key events that have affected it (delays, technical problems, organizational problems,...) through a non-participant observation that consisted in attending important meetings organized by the project manager (team meeting, steering committee) and exchanging informally with the members of the project team, the contributors to the project belonging to the Architecture unit, or other project stakeholders. The approach adopted to carry out our case study is an inductive qualitative method where field data are used to highlight concepts representative of the studied phenomenon (Thiétart, 1999).

Analysis of the information collected has confirmed the interest of this case study both at the level of the theoretical model evaluation, and in terms of its managerial implications. First, the organization of Information Systems Department of the the insurance group French subsidiary studied in this paper has many similarities with the organization suggested by the computerization project model presented above. Indeed, apart from the project management space represented by the project manager, the Architecture unit can be likened to the solution space while the construction space consists of the Development unit and the Indian "offshore" company. The problem space is composed of representatives of the NREP customers who are either members of the marketing department, or financial advisors. End users of NREP tool are grouped in the operation space. It follows that the computerization project model is appropriate to model the Information Systems Department which carries out the development of the NREP tool within the insurance group. Second, information collected allowed us to verify the existence of the knowledge boundaries identified by the theoretical model proposed in this paper.

On the one hand, the fact that the definition of the users needs and requirements is not described formally confirms the existence of a knowledge boundary between the problem side (problem space, operation space) and the solution side (solution space, construction space). This boundary is the source of users requirements definition problems which result in the failures of many computerization projects. The solutions proposed in the literature to manage this boundary was not fully effective as they have analyzed this boundary holistically. In accordance with Brooks analysis (Brooks, 1987), our model considers that that the knowledge boundary between the problem and the solution sides may be broken down into four types: Customer – Architect, Project manager - Customer, Project manager - End user, Developer - End user. This decomposition permits suggesting appropriate means to handle each type knowledge boundary.

On the other hand, the diversity of professions within the solution side confirms the existence of knowledge boundaries both between the Architecture and Development units and within each unit. Indeed, the Architecture unit is peopled of experts who do not use the same practices and the same professional concepts while performing different activities. For example, the concepts and practices used by functional architects and database experts are different. Similarly designers and developers do

not always understand the usefulness of architectural rules and methodological and quality assurance standards, imposed by the Architecture unit.

Furthermore, within the problem side, the diversity of end users and their representatives shows the existence of knowledge boundaries between end users and their representatives, between end-users, or between representatives of end users. For example, the concepts, practices, and interests of financial advisors differ significantly from those used by the marketing department members. Therefore, our model facilitates the detailed identification of knowledge boundaries in computerization projects. As a result, the application of this model to information collected as part of our case study permit us verifying the existence of the knowledge boundaries listed above (Table 1). In the rest of this section, we have focused on a subset of Knowledge borders whose analysis summarizes the main findings of our case study. These are the following knowledge boundaries.

In the remainder of this section, we have focused on a subset of knowledge boundaries whose analysis summarizes the main conclusions of our case study. These knowledge boundaries are the following:

- a knowledge boundary between the customer representatives and the functional architect involved in the project,
- a knowledge boundary between the functional architect and the technical architect involved in the project,
- a knowledge boundary between the functional and technical architects and the designers,
- a knowledge boundary between the designers and the offshore programmers,
- a knowledge boundary between the methods engineers and the project team,

Knowledge boundaries management took place through a series of decisions taken by the project manager and validated by the governance authorities. These decisions have impacted both the project team that the forms of collaboration between the various project stakeholders. Thus, a facilitator intervened during the early stages of the project to help the functional architect to understand all the business aspects to be taken into account in defining the architecture of the NREP tool. This facilitator is a former financial advisor who joined in recent years Information Systems Department as a manager of the computer systems technical and functional problems. This decision helps manage the knowledge boundary between the customer representatives and the functional architect involved in the project. Based on the information we have collected, the facilitator was found to be effective by the actors involved in this knowledge boundary between the customer representatives and the functional architect involved in the project. Therefore, the facilitator can be considered as a clarification boundary object which permits managing the Customer – Architect inter-spaces boundary. However, a business dictionary uniquely defining the different business concepts can also be considered as an additional clarification boundary object complementary to the facilitator boundary object.

The knowledge boundary between the functional architect and the technical architect involved in the project is an intra-space boundary that materializes a misunderstanding between these two actors due to the lack of a shared professional culture. Indeed, the technical architect is a former programmer who does not master the software application architecture (components, connectors) and has only vague notions of the functional architecture and business intelligence architecture. The functional architect has no significant experience in defining the software architecture of an application and does not master all the constraints related to the interfaces between components or between applications. To prevent these divergences of views between the technical and the functional architect, the project manager suggested that the functional architect introduces the technical architect to the basic concepts of decisional architecture. Although this training is deemed positive by the technical architect, his divergences with the technical architect persisted throughout our observation period. These divergences of views were compounded by a lack of trust between the two actors. They often surfaced in the meetings of the project team and working sessions with designers. Despite these viewpoints differences, the software architecture of the NREP tool has been developed on the basis of the

functional architecture. Also, the training provided by the functional architect played only partially the role of clarification boundary object

The knowledge boundary between the architects (functional and technical) and the designers is an inter-spaces boundary that has been managed in three stages. First, a training session was conducted by the functional architect and the technical architect in order to familiarize the two designers with the basic concepts and constraints of the information systems architecture, and convince them that these concepts are useful. This training can be considered as a clarification boundary object unlike architectural guides, published by the Architecture unit, which was considered incomprehensible by the two designers. Second, prototypes of the NREP tool functional and software architectures were developed iteratively by both architects and discussed by the two designers over several working sessions. The functional and software architectures NREP tool have been defined following these sessions. Third, it was decided to immerse part-time a technical architect in the development teams (design, coding) to help designers and programmers implement the NREP tool functional and software architectures and verify that all features of this solution was taken into account. Apart from training, four boundary objects can be identified at this stage. There are two clarification boundary objects and two objects production boundary objects. The NREP tool functional and software architectures prototypes and the technical architect associated with the development team are clarification boundary objects. The NREP tool functional and software architectures are both production boundary objects.

The knowledge boundary between designers and offshore programmers is a complex intra-spaces boundary difficult to manage. The complexity of this boundary is due to the existence of three types of dispersions: geographic, linguistic, and cultural. The geographical dispersion reflects the remoteness of design and implementation teams. This remoteness was compounded by the linguistic dispersion caused by the problems of spoken and written languages of programmers and designers. Thus, the communication between the two parties - which took place by telephone, mail, or video conferencing - has been difficult and resulted in many misunderstandings. The cultural dispersion, caused by differences in programmers and designers national cultures has fostered a climate of mistrust between the two parties which constituted a barrier to knowledge sharing. The communication difficulties between the programmers and designers, and the lack of English translation of documents and design standards to be respected explain the fact that there was no boundary object contributing the knowledge boundary management. In order to cope with the risks of failure of the project, the project manager has proposed to call upon an international consulting company composed of information systems consultants mastering both the English and French languages, including Indian consultants. Following the approval of the project manager decision by the governance authorities, the international consulting company became the interlocutor of the project team and the architects and the mediator between them and the programmers of the Indian offshore company. This company has performed the English translation of the documents provided by the designers and the standards to be applied, and reported on the progress of programming work to the project manager. Similarly, it has realized the French translation of the documents prepared by the Indian programmers and provided a French version of the NREP tool to the Information Systems Departement for integration tests. At this stage, many boundary objects can be identified. For example, the French version of the NREP tool, the implementation documents translated into French, and the design documents translated into English are transfer boundary objects. We note that among these boundary objects, some will be packaged and delivered to end users. In this case, they will become production boundary objects. The status and progress reports provided by the international consulting company are communication boundary objects. Finally, the international consulting company can be considered as a clarification boundary object.

The knowledge boundary between methods engineers and the project team reflects the resistance of the designers to the application of the development defined and recommended by the Architecture unit. Although this method enables software prototyping and iterative development, it is perceived by designers as being too restrictive and a source of inconsistencies and long delays. To circumvent the designers resistance, the project manager has proposed to the methods engineers to cooperate with the

designers in order to identify the methodological aspects and documentary standards whose application will have a significant added value for the project progress. This decision has not been approved by the governance authorities who feel that its implementation will be costly and cause delays. Thus, no boundary object was built to manage this knowledge boundary.

This case study presents three advantages. On the one hand, it allowed us to verify the interest of the approach proposed in this paper to identify the knowledge boundaries in a computerization project, and analyze the boundary objects that help manage them. Indeed, the existing studies of computerization projects knowledge boundaries are often descriptive and do not provide sufficient information to facilitate the knowledge sharing and help the project governance authorities take the appropriate decisions. On the other hand, it has provided information confirming the need to clarify the concept of boundary object. Since boundary objects are intended to facilitate knowledge sharing, such objects can take many forms. They can be informational artifacts, human beings with multiple skills, or organizational entities. This confirms some of the results obtained by Levina and Vaast (2005) and Carton and Farastier (2012) who have introduced the boundary actor concept. Moreover, to be effective, boundary objects in the form of informational artefacts must be built collectively by organizational actors involved in knowledge boundaries. This explains the positive role played by the training sessions, the informational prototypes in the management of knowledge boundaries in our case study. This also explains why the guides and standards built by a group of actors are not always effective boundary objects. Finally, we use this case study to highlight the following managerial implications:

- To manage inter-spaces knowledge boundaries, one can effectively use human actors with multiple skills as boundary objects.
- To manage the knowledge boundaries resulting from differences in national cultures or spoken or written languages, organizational entities can be effective boundary objects.
- To manage a knowledge boundary with a boundary object in the form of an informational artifact, all organizational actors involved in this boundary must participate in the construction of this boundary object.
- Setting up a business dictionary can help manage knowledge boundaries between the problem side and the solution side.

7 Conclusion and future reserach directions

In this paper, we proposed a theoretical framework to analyze the knowledge boundaries in a computerization project and identifying the boundary objects that help manage them. The computerization project model used in this work is a generic computerized approach and not a particular computerization method. It is therefore open to all computeraization approaches and methods whatever their nature. The identification of computerization projects knowledge boundaries and boundary objects that help manage them is the first contribution of our article that provide an answer to the first research question. The interest of this analysis is related on the one hand, to its independence of existing computerized methods, and on the other hand, in the consideration - in each identified knowledge boundary - of all organizational actors who participate to it. For example, the Customer – Architect knowledge boundary takes into account not only the customer and the architect types but also the developer and end user types who play a secondary role respectively in solution space and the problem space. That each type of computerization project knowledge boundary includes many heterogeneous organizational actors who hold specialized knowledge permits highlighting the complexity of such boundaries and the difficulty to manage them. In addition, we have answered the second research question by defining a typology of boundary objects in the context of a computerization project and providing examples of such objects. The second contribution of our paper is related to the accuracy of the boundary object concept. Indeed, we have established that a boundary object can be either an informational artifact, an organizational actor, or an organizational entity. The third contribution of our paper results from the managerial implications of the approach we have

proposed for the knowledge boundaries management. However, the analysis in this article should be more detailed and enough thorough to describe the production boundary objects issued from the political approach. This is to specify in particular the role of national and organizational cultures in the shaping of production boundary objects if innovation and instability are not limited to technology but include social and cultural factors. This is the case, for example, of computerization projects in virtual organizations or post-merger organizations. Another important direction of research is to use the findings of this paper to build instruments that permit assessing the ability of computerization methods and approaches to promote the knowledge sharing.. Finally, the findings from the case study presented in this paper should be enhanced and consolidated with other case studies. This is a third research direction.

References

- Adler, P. and S. W. Kwon (2002). "Social capital: Prospects for a new concept." *Academy of Management Review* 27 (1), 17-40.
- Alchian, A. A. and H. Demsetz (1972). "Production, Information Costs and Economic Organization." *American Economic Review* 62 (5), 777-795.
- Argote, L. (1999). *Organizational Learning: Creating, Retaining, and Transferring Knowledge*. Norwell, MA: Kluwer.
- Arthur, M., De Fillippi, R. and C. Jones (2001). "Project-based learning as the interplay of career and company nonfinancial capital." *Management Learning* 32 (1), 97-11.
- Baetjer, H., Jr. (1998). "Software as Capital: An Economic Perspective on Software Engineering." Piscataway, New Jersey: The Institute of Electrical and Electronics Engineers.
- Becker, M. C. (2001). "Managing dispersed knowledge: organizational problems, managerial strategies, and their effectiveness." *Journal of Management Studies* 38 (7), 1037-1051.
- Ben Chouikha M. and S. Ben Dhaou Dakhli (2011). "Knowledge Sharing Enablers and Barriers: The Case of Virtual Organizations." In: *Proceedings of the IFIP-AIS MCIS'2011 conference*, Limassol, Cyprus, September 3-5, 2011.
- Ben Chouikha, M. (2010). "Design de l'Organisation Apprenante : Etude de Cas Longitudinale des Equipes de Projet NOKIA SIEMENS NETWORKS. " Thèse de doctorat, Université de Paris-IX Dauphine, Paris.
- Birnholtz, J. and T. Finholt (2007). "Cultural Challenges to Leadership in Cyberinfrastructure Development." In: *Leadership at a Distance: Research in Technology- Supported Work*, S. Weisband (ed.), New York: Lawrence Erlbaum Associates, 195-207.
- Boehm B. W. (1988). "A Spiral Model of Software Development and Enhancement." *IEEE Computer* 21 (5), 61-72.
- Boehm, B. W. (1976). "Software Engineering." *IEEE Transactions on Computers* C25 (12), 1226-1241.
- Boehm, B.W. (1984). "Verifying and Validating Software Requirements and Design specifications." *IEEE Software* 1 (1), 75-88.
- Bourdieu, P. (1984). *Distinction: A Social Critique of the Judgment of Taste*. Cambridge, MA: Harvard University Press.
- Bredillet, C. N. (2004). "Projects: Learning at the edge of organization." In: P. W. Morris and J. K. Pinto (Eds.), *The Wiley Guide to Managing Projects*. Hoboken: John Wiley & Sons.
- Briers M. and W. F. Chua (2001). "The role of actor-networks and boundary objects in management accounting change: a field study of an implementation of activity-based costing." *Accounting, Organizations and Society* 26, 237-269.
- Brooks, F. P. Jr. (1987). "No Silver Bullet-Essence and Accidents of Software Engineering." *IEEE Computer* 20 (4), 10-19.
- Brooks, F. P. Jr. (1995). *The Mythical Man Month: Essays on Software Engineering*. Reading, M.A.: Addison-Wesley.

- Brown, J. S. and P. Duguid (2001). "Knowledge and organization: A socialpractice perspective." *Organization Science* 12, 198–213.
- Carlile, P. R. (2002). "A Pragmatic View of Knowledge and Boundaries: Boundary Objects in New Product Development." *Organization Science*, 13 (4), 442-455.
- Carlile, P. R., and C. Østerlund (2005). "Relations in Practice: Sorting Through Practice Theories on Knowledge Sharing in Complex Organizations." *Information Society* 21 (2), 91-107.
- Carlile, P. R. and E. S. Rebentisch E. S. (2003). "Into the Black Box: The Knowledge Transformation Cycle." *Management Science* 49 (9), 1180-1195.
- Carlile, P. R. (2004). "Transferring, Translating, and Transforming: An Integrative Framework for Managing Knowledge across Boundaries." *Organization Science* 15 (5), 555-568.
- Carton, S. and A. Farastier (2012). "Intégration des connaissances client dans un projet en systèmes d'information : influence de l'environnement de connaissance du projet." *Revue Systèmes d'Information et Management* 17 (2).
- Chan, R. and M. Rosemann (2001). "Managing knowledge in enterprise systems." *Journal of Systems and Information Systems* 5 (2), 37–53.
- Chan, Y. E. (2001). "Why Haven't We Mastered Alignment? The Importance of the Informal Organization Structure." *MIS Quarterly Executive* 1 (2), 97-112.
- Coase, R. (1937). "The Nature Of The Firm." *Economica* 4, 386-405, Traduction Française: *Revue Française d'Economie* 2 (1), 133-163.
- Coupe, R. T. and N. M. Onodu (1996). "An empirical evaluation of the impact of CASE on developer productivity and softw are quality." *Journal of Information Technology* 11 (2), 173-181.
- Cramton, C. D. and P. J. Hinds (2007). "Intercultural Interaction in Distributed Teams: Saliency of and Adaptations to Cultural Differences." In: *Proceedings of the Academy of Management Annual Meeting, Best Papers*. G. Salomon (ed.), Philadelphia, PA, August 3-8.
- Cramton, C. D. (2001). "The Mutual Knowledge Problem and Its Consequences for Dispersed Collaboration." *Organization Science* 12 (3), 346-371.
- Dakhli, S. (1998). "Le Prototypage." Thèse de doctorat, Université de Paris-IX Dauphine, Paris.
- De Vaujany, F.-X. (2009). *Les grandes approches théoriques du système d'information*. Paris Editions Hermès-Lavoisier.
- DiMaggio, P. J. (1991). "Social Structure, Institutions, and Cultural Goods: The Case of the United States." In: *Social Theory for a Changing Society*. P. Bourdieu and J. S. Coleman (eds.), Boulder, NY: Westview Press, pp. 133-155.
- Ekstedt, E. (2002). "Contracts of work in a project based economy." in Sahlin Andersson, K., Söderholm, A. (Eds.), *Beyond project management – New perspectives on the temporary – permanent dilemma*. Liber, Malmö: pp. 59-77.
- Engwall, M. (2003). "No project is an island: linking projects to history and context, Research Policy." 32 (5), 789-808.
- Espinosa, J. A., Cummings, J. N., Wilson, J. M. and B. M. Pearce (2003). "Team Boundary Issues across Multiple Global Firms." *Journal of Management Information Systems*. 19 (4), 157-190.
- Faraj, S. and L. Sproull (2000). "Coordinating expertise in software development teams." *Management Science* 46 (12), 1554–1568.
- Fayad, M., Henry D. and D. Bougali (2002). "Enterprises Frameworks." *Software Practice and Experience* 32, 735-786.
- Feng, Y., Ye, H. and S. L Pan (2011). "Delivering Knowledge Across Boundaries: A Case Study of Bankco's Offshoring Projects." *Pacific Asia Journal of the Association for Information Systems* 3 (3), 57-79. Available at: <http://aisel.aisnet.org/pajais/vol3/iss3/4>.
- Finlay, P. N. and A. C. Mitchell (1994). "Perceptions of the Benefits From the introduction of CASE: An Empirical Study." *MIS Quarterly* 18 (4), 353-370.
- Gurbaxani V. and S. Whang (1991). "The Impact of Information Systems on Organizations and Markets." *Communications of the ACM* 34 (1), 59-73.

- Hardy, C., Lawrence, T. B. and D. Grant (2005). "Discourse and Collaboration: The Role of Conversations and Collective Identity." *Academy of Management Review* 30 (1), 58-77.
- Henderson, J. C. and N. Venkatraman (1994). "Strategic Alignment: A Model for Organizational Transformation via Information Technology." In: Allen, T. J. and Scott Morton, M. S. (Eds.), *Information Technology and the corporation of the 1990s*. Research Studies. Oxford: Oxford University Press.
- Henderson, J. C. and N. Venkatraman (1993). "Strategic alignment: Leveraging information technology for transforming organizations." *IBM Systems Journal* 32 (1), 472-484.
- Hinds, P. J. and M. Mortensen (2004). "Understanding Conflict in Geographically Distributed Teams: The Moderating Effects of Shared Identity, Shared Context, and Spontaneous Communication." *Organization Science* 16 (3), 290-307.
- Hirschheim, R. and R. Sabherwal (2001). "Detours toward Strategic Information Systems Alignment." *California Management Review* 44 (1), 87-108.
- Hoegl, M. and H. G. Gemuenden (2001). "Teamwork Quality and the Success of Innovative Projects: A Theoretical Concept and Empirical Evidence." *Organization Science* 12 (4), 435-449.
- Horio, H. and C. Watanabe (2008). "The Paradox of a Service-Oriented Economy for Sustainability: Co-Evolution between Innovation and Resources Effectuation by a Global Complement." *Journal of Services Research* 8 (1), 155-176.
- Jensen, M. C. and W. H. Meckling (1976). "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3 (4), 305-360.
- Kakati, M. (2002). "Mass customization – needs to go beyond Technology." *Human Systems Management* 21, 85-93.
- Kaulio, M. A. (2008). "Project leadership in multi-project settings: Findings from a critical incident study." *International Journal of Project Management* 26 (4), 338-347.
- Koskinen, K. U., Pihlanto, P. and H. Vanharanta (2003). "Tacit Knowledge Acquisition and Sharing in a Project Work Context." *International Journal of Project Management* 21 (4), 281-290.
- Kotha, S. (1995). "Mass customization: Implementing the Emerging Paradigm for Competitive Advantage." *Strategic Management Journal* 16, Special Issue, 21-42.
- Krishna, S., Sahay, S. and G. Walsham (2004). "Cross-Cultural Issues in Global Software Outsourcing." *Communications of the ACM* 47 (4), 62-66.
- Lam, A. (1997). "Embedded Firms, Embedded Knowledge: Problems of Collaboration and Knowledge Transfer in Global Cooperative Ventures." *Organization Studies* 18 (6), 973-996.
- Leavitt, H. J., (ed.) (1963). *The Social Science of Organizations, Four Perspectives*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Leonard-Barton, D. (1995). *Well Springs of Knowledge: Building and Sustaining the Sources of Innovation*. Boston, MA: Harvard Business School Press.
- Leseure, M. J., and N. J. Brookes (2004). "Knowledge management benchmarks for project management." *Journal of Knowledge Management* 8 (1), 103-116.
- Levina, N., and E. Vaast (2008). "Innovating or Doing as Told? Status Differences and Overlapping Boundaries in Offshore Collaboration." *MIS Quarterly* 32 (2), 307-332.
- Levina, N., and E. Vaast (2005). "The Emergence of Boundary Spanning Competence in Practice: Implications for Implementation and Use of Information Systems." *MIS Quarterly* 29 (2), 335-363.
- Levina, N., and E. Vaast (2006). "Turning a Community into a Market: A Practice Perspective on It Use in Boundary-Spanning." *Journal of Management Information Systems* 22 (4), 13-38.
- Levina, N. (2005). "Collaborating on Multiparty Information Systems Development Projects: A Collective Reflection-in-Action View." *Information Systems Research* 16 (2), 109-130.
- Litwak, E. and L. F. Hylton (1962). "Interorganizational Analysis: A hypothesis on coordination agencies." *Administrative Science Quarterly* 6, 395-420.
- Longépé, C. (2006). *Le Projet d'Urbanisation du SI*. Paris: Editions Dunod.
- Lundin, R. A. and A. Soderholm (1995). "A Theory of the temporary organization." *Scandinavian Journal of Management* 11 (4), 437-455.

- Maaninen-Olsson, E., Wismén, M. and A. A. Carlsson (2008). "Permanent and temporary work practices: knowledge integration and the meaning of boundary activities." *Knowledge Management Research & Practice* 6 (4), 260-273.
- Manning, S. (2008). "Embedding projects in multiple contexts – a structuration perspective." *International Journal of Project Management* 26 (1), 30-37.
- Meehan, B. and I. Richardson (2002). "Identification of software process knowledge management." *Software Process Improvement and Practice* 7 (2), 47–55.
- Metiu, A. (2006). "Owning the Code: Status Closure in Distributed Groups." *Organization Science* 17 (4), 418-435.
- Montgomery, K. and A. L. Oliver (2007). "A Fresh Look at How Professions Take Shape: Dual-Directed Networking Dynamics and Social Boundaries." *Organization Studies* 28, 1-27.
- Newell, S., Tansley, C. and J. Huang (2004). "Social Capital and Knowledge Integration in an ERP Project Team: The Importance of Bridging AND Bonding." *British Journal of Management* 15, 43-57.
- Nonaka, I. and I. Takeuchi (1995). *The Knowledge-Creating Organization*. Oxford: Oxford Press.
- Packendorff, J. (1995). "Inquiring into the temporary organization: New directions for project management research." *Scandinavian Journal of Management* 11 (4), 319-333.
- Putnam, R. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon and Shuster.
- Reich, B. H. and S. Y. Wee (2006). "Searching for knowledge in the PMBOK® guide." *Project Management Journal* 37 (2), 11-25.
- Reich, B. H., Gemino, A. and C. Sauer (2008). "Modeling the Knowledge Perspective of IT Projects." *Project Management Journal* 39, Supplement, S4–S14.
- Reich, B. H. (2007). "Managing Knowledge and Learning in IT Projects: A Conceptual Framework and Guidelines for Practice." *Project Management Journal* 38 (2), 5-17.
- Ross, J. W., Weill, P. and D. Robertson (2006). *Enterprise architecture as strategy: Creating a Foundation for Business Execution*. Boston, MA: Harvard Business Press.
- Sabherwal, R. and Y. E. Chan (2001). "Alignment Between Business and IS Strategies: A Study of Prospectors, Analyzers, and Defenders." *Information Systems Research* 12 (1), 11-33.
- Star, S. and J. Griesemer (1989). "Institutional Ecology, 'Translations', and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-1939." *Social Studies of Science* 19 (3), 387-420.
- Thiéart, R. A. (1999). "Méthodes de recherche en management." Paris: Editions Dunod.
- Toffolon, C. and S. Dakhli (2002). "The Software Engineering Global Model." In: *Proceedings of the COMPSAC'2002 Conference*, Oxford, United Kingdom. Los Alamitos, CA: IEEE Computer Society Press, (pp. 47-53)
- Toffolon, C. (1996). "L'Incidence du Prototypage dans une Démarche d'Informatisation." Thèse de doctorat, Université de Paris-IX Dauphine, Paris.
- Toffolon, C. (1999). "The Software Dimensions Theory." In: Joaquim Filipe (Ed.), *Enterprise Information Systems, Selected Papers Book*, KLUWER ACADEMIC PUBLISHERS.
- Vinck, D. (2009). "De l'objet intermédiaire à l'objet-frontière: Vers la prise en compte du travail d'équipement." *Revue d'anthropologie des connaissances* 3 (1), 51-72.
- Walsham, G. (2002). "Cross-Cultural Software Production and Use: A Structural Analysis." *MIS Quarterly* 26 (4), 359-380.
- Walz, D. B., Elam, J. J. and B. Curtis (1993). "Inside a software design team: Knowledge acquisition, sharing and integration." *Communications of the ACM* 36 (10), 63–77.
- Williamson, O. E. (1981). "The Modern Corporation: Origins, Evolution, Attributes." *Journal of Economic Literature* 19, 1537-1568.
- Winter, S. and G. Szulanski (2001). "Replication as strategy." *Organization Science* 16, 730–743.
- Yin, R. K. (1994). *Case Study Research: Design and Methods*. Beverly Hills, CA: Sage Publishing.

DID THE REDUCTION OF ICT INVESTMENT DUE TO THE 2008 ECONOMIC CRISIS AFFECT THE INNOVATION PERFORMANCE OF FIRMS? – AN EXPLORATORY ANALYSIS BASED ON FIRM DATA FOR THE EUROPEAN GLASS, CERAMICS AND CEMENT INDUSTRY

Completed Research

Arvanitis, Spyros, ETH Zurich, Zurich, Switzerland, arvanitis@kof.ethz.ch

Loukis, Euripidis, University of Aegean, Samos, Greece, eloukis@aegean.gr

Abstract

In this paper we investigate, first, the characteristics of the firms that reduced their ICT investment due to the 2008 crisis, focusing on the firms' ICT characteristics in terms of the ICT, budget, skills and applications used. The analysis of the ICT characteristics that may influence the likelihood of having reduced ICT investment as a consequence of the crisis is primarily explorative, thus driven by available data and economic intuition. The second research question we examine empirically refers to the possibility that an economic crisis could affect innovation performance through the ICT investment channel. In connection with this, it is also interesting to analyze the ICT characteristics that are associated with ICT-enabled innovation performance; this is the third research question of this paper. Our study is based on firm data from the glass/ceramics/cement industry in six European countries, which have been collected through the 'e-Business Survey' of the European Union. We find that ICT-related crisis vulnerability correlates positively with decreasing ICT budgets (pro-cyclical investment behaviour), the existence of skill deficits in ICT, the awareness of and interest in new ICT applications that presumably request much additional ICT investment, the exposure to strong price competition and the strong presence in international markets, in which activities have significantly decreased due to the crisis. Further, statistically significant negative relationship between ICT-enabled product innovation and crisis vulnerability (pro-cyclical behaviour) is found only for new products or services that contain ICT components, and are therefore directly affected by crisis-related decreasing product demand.

Keywords: economic crisis, ICT, innovation, ICT-enabled innovation

1 Introduction

The financial and economic crisis of 2008 affected negatively investment in general and also investment in ICT was not left unchanged (Rojko et al. 2011, OECD 2009). World ICT spending fell by around 4% in 2009 (OECD 2010). Nevertheless, the decrease of ICT investment has been lower than that of GDP worldwide so that the ratio of ICT investment to GDP has increased. The 2009 decline of world spending in ICT is not as large in current US dollars as in 2001-02, owing to growth in non-OECD economies and the introduction of new products (OECD 2010). Worldwide, about 57% of the 2009 spending was on communication services and hardware, 21% on computer services, 13% on computer hardware and 9% on software. About 76% of world ICT spending is in the OECD countries (OECD 2010).

The expectation for the impact of the crisis on ICT investment is qualitatively the same as for all other kinds of investment. The main idea is that independent of the source of financing the general investment propensity decreases in periods of economic recession. Firms are confronted with demand uncertainty that makes investment more risky than in 'normal' or boom periods. This demand uncertainty forces firms to a pro-cyclical behaviour. Particularly in the innovation literature, an alternative approach is discussed that leads to an anti-cyclical investment behavior. According to the opportunity costs approach in a booming economy it is expected that costs for labour and other input factors for R&D activities are high and in recessions costs are clearly lower. Hence, opportunity costs are lower in recessions and firms would benefit if they could shift resources to R&D activities. It is then an empirical issue which situation for which investment type prevails.

We are interested to investigate, first, the characteristics of the firms that reduced their ICT investment due to the 2008 crisis, particularly the firms' ICT characteristics in terms of the ICT budget, skills and applications used. The analysis of the ICT characteristics that may influence the likelihood of a reduction of ICT investment as a consequence of the crisis is primarily explorative, thus driven by available data and economic intuition.

ICT use is according to existing empirical literature an important driver of innovation activities and performance (Kleis et al. 2012). As a consequence, we expect that the drop of ICT investment during the crisis would cause a negative effect on specific ICT-enabled innovation performance. This would be the ICT-related effect of crisis on innovation. Thus, the second research question we want to examine empirically refers to the possibility that an economic crisis could affect innovation performance through the ICT investment channel (i.e. by leading to lower ICT investment). In connection with this, it is also interesting to analyze the ICT characteristics that are associated with ICT-enabled innovation performance. This is the third research question we want to pursue in this paper.

Our study is based on firm data from the glass/ceramics/cement industry in six European countries, which have been collected through the 'e-Business Survey' of the European Union. The glass, ceramics and cement industry is an important supplier of the construction sector, and as a consequence, an industry that is dependent primarily on domestic demand. In this sense, the state of affairs in this industry reflects the situation of a country's part of the economy that is oriented to domestic demand.

We find that ICT-related crisis vulnerability correlates positively with decreasing ICT budgets (pro-cyclical investment behaviour), the existence of skill deficits in ICT, the awareness of and interest in new ICT applications that presumably request much additional ICT investment, the exposure to strong price competition and the strong presence in international markets, in which activities have significantly decreased due to the crisis. Further, statistically significant negative relationship between ICT-enabled product innovation and crisis vulnerability (pro-cyclical behaviour) is found only for new products or services that contain ICT components, and are therefore directly affected by crisis-related decreasing product demand. To our knowledge, there is no other study investigating these topics, so our paper has also the character of an explorative study in a new research field.

The paper is structured as follows: section 2 presents the conceptual background, related literature and the research hypotheses. Section 3 discusses the data. In section 4 is our methodology - models specification presented, and also econometric issues are discussed. In section 5 the results are presented and discussed. Section 6 concludes the paper.

2 Conceptual background, related literature and research hypotheses

2.1 ICT Investment, Innovation Investment and Crisis

The expectation for the impact on ICT investment is qualitatively the same as for all other kinds of investment. The main idea is that independent of the source of financing the general investment propensity decreases in periods of economic recession. Firms are confronted with demand uncertainty that makes investment more risky than in 'normal' or boom periods. Decreasing demand limits also internal financing of investment by past revenues. Uncertain economic perspectives reduce also the willingness of banks and other financial intermediaries to finance firms' investment projects. Of course, not all kinds of investment bear the same risk, innovation projects being considered as quite risky, buildings being seen as much less risky than other investment categories (see, e.g., Kahle and Stulz 2010; Gerner and Stegmaier 2013; Geroski and Gregg 1997). Further, not all types of firms bear the same risk, small firms being confronted with more difficulties to finance investments in recession than large firms due to credit rationing by financial intermediaries (for the theoretical background see, e.g., Stiglitz and Weiss 1981 for investment in general; Goodacre and Tonks 1995 for investment in innovation). In general, we expect that economic crisis negatively affects ICT investment. To our knowledge, there are no studies dealing with question of the impact of economic crisis on ICT investment.¹ As a consequence, it is not a priori clear which ICT characteristics explain a firm's behavior in an economic crisis (see hypothesis 1a to 1c below). The analysis of the ICT characteristics that may influence the likelihood of having reduced ICT investment as a consequence of the crisis is primarily explorative, thus driven by available data and economic intuition.

There is some theoretical consent (see, e.g., Barlevy 2007) and some empirical evidence (see, e.g., Quyang 2011a and Guellec and Ioannidis 1999) that R&D investment expenditures of firms, the most important input for innovation, are pro-cyclical, i.e. they are increasing in the business upswing and they are decreasing in the business downturn. However, there are also some theoretical arguments as well as some anecdotal evidence that firms show an anti-cyclical R&D investment behaviour. In order to explain pro- or anti-cyclical R&D behaviour we have to take into account two diverging forces, i.e. demand aspects (see Filippetti and Archibugi 2011) and the opportunity costs aspect (Rafferty and Funk 2004).² Since R&D investments are predominantly financed through the cash-flow of a firm, which is expected to fluctuate pro-cyclically with the demand, we would expect a pro-cyclical R&D investment behaviour as well. In a booming economy it is expected that costs for labour and other input factors for R&D activities are high and in recessions costs are clearly lower. Hence, opportunity costs are lower in recessions and firms would benefit if they could shift resources to R&D activities. If the opportunity cost aspect prevails, we would expect anticyclical behaviour; firms would make use of lower production costs in recessions and would intensify their R&D investments. Since empirical evi-

¹ However, there is a paper which is worth mentioning: In a case study, Leidner et al. (2003) found based on interviews with 20 CIOs that firms reacted both pro- and anti-cyclically to the crisis of 2000-02 depending on their short-term or long-term perspective.

² For a more detailed discussion of these approaches see Arvanitis and Woerter (2014).

dence supports mostly pro-cyclical behaviour we concentrate here on the hypothesis of pro-cyclical behaviour with respect to innovation (see hypothesis 2 below).

2.2 ICT and Innovation

There are many ways through which the use of ICT can contribute to firms' innovation. According to Kleis et al. (2012) this can happen through three main channels. The first channel goes through the improvement of the management of the knowledge used in the innovation process. Information technology enables an efficient storage and a high accessibility of this knowledge throughout an enterprise. Internal networks, e-mail systems, and electronic databases all facilitate the transfer of knowledge and the communication between innovation participants inside the firm. Second, ICT enables a more efficient cooperation in innovation with external partners. Information technology facilitates the exchange of information with external partners that are located far away from the focal firm. The creation of new knowledge through collaboration with other firms has become more and more important in the last twenty years (Enkel et al. 2009). Third, ICT contributes directly to the innovation production in several ways. Kleis et al. (2012) identified three main stages of the innovation process, for which the application of ICT has proved to be useful. First, the stage of the generation of ideas for new products can benefit from information systems (e.g., Customer Relationship Management CRM) that enable a firm to analyse customers' communication and transaction data, and identify needs that can be covered by new products or significant modifications of existing products. Second, ICT enables the development of efficient design capabilities for new products. For example, technologies such as computer-aided design (CAD) and computer-aided manufacturing (CAM) help to digitize a new product's design, make it available throughout the innovation process. Third, ICT helps integrate design and production systems, so that errors of information transfer and translation are reduced and, as a consequence, the efficiency of this last stage of the innovation process is increased (Tapscot et al., 2000; Brynjolfsson and Saunders, 2010). In sum, we expect a positive impact of ICT through these three channels on innovation performance (see hypothesis 3 below).

2.3 Research Hypotheses

Based on the above discussion of literature, we postulate a series of hypotheses:

Hypothesis 1a: The ICT characteristics that refer to ICT inputs (budget, human capital) and the ICT infrastructure (as represented by the ICT applications used by a firm) are jointly positively correlated with the ICT-related crisis vulnerability.

Hypothesis 1b: The ICT characteristics that refer to ICT inputs (budget, human capital) and the ICT infrastructure (as represented by the ICT applications used by a firm) are jointly negatively correlated with the ICT-related crisis vulnerability.

Hypothesis 1c: The ICT characteristics that refer to ICT inputs (budget, human capital) and the ICT infrastructure (as represented by the ICT applications used by a firm) are not at all correlated with the ICT-related crisis vulnerability.

Hypothesis 2: ICT-enabled innovation performance measures are negatively correlated with the extent of a firm's vulnerability with respect to ICT-related crisis. Thus, we expect pro-cyclical behaviour with respect to ICT-enabled innovation.

Hypothesis 3: The ICT characteristics that refer to ICT inputs (budget, human capital) and the ICT infrastructure (as represented by the ICT applications used by a firm) are jointly positively associated with ICT-enabled innovation performance measures.

3 Data

The data come from the e-Business Survey 2009 of the European Union. This survey among companies from the glass, ceramics and cement industries consisted of 676 telephone interviews with ICT decision-makers in six EU countries: Germany, Spain, France, Italy, United Kingdom and Poland. Interviews were carried out in March 2009, using computer-aided telephone interview (CATI) technology. The questionnaire contained about 90 questions which were structured into eight modules. The survey population was defined as companies with at least 10 employees which used computers, were active within the national territory of one of the six countries covered, and which had their primary business activity in the glass, ceramics or cement industry as specified by NACE Rev. 2 Groups 23.1-6. The survey was carried out as an enterprise survey: data collection and reporting focus on the enterprise, defined as a business organization (legal unit) with one or more establishments. The sample drawn was a stratified random sample of companies from the population in each of the six countries, with the objective of fulfilling minimum strata with respect to company size-bands per country-sector cell (see 'Annex I: Methodology Report' in Sectoral e-Business Watch 2009). Pilot interviews prior to the regular fieldwork were conducted with about 15 companies in Germany in February 2009, in order to test the questionnaire (structure, comprehensibility of questions, average interview length). The response rate, i.e. the number of completed interviews divided by the net sample of contacts established with eligible hospitals/enterprises, was typically about 15-20%, with, however, big differences in some of the countries. The dataset that was used for the econometric estimates contained 676 observations.

Table 1 contains information on the impact of economic crisis on ICT investment plans or projects by subsector, country and firm size class. Crisis vulnerability has been rather uniformly distributed among the firms of the three subsectors: about 41% (glass) to 48% (ceramics) firms reported an impact of crisis on their ICT investment plants. Small firms have been less crisis vulnerable than middle-sized and large firms, presumably because larger firms are stronger dependent than small firms on exports that have decreased significantly as a consequence of the crisis. Significant differences exist among countries: about 26% reported a crisis impact in Germany, while the respective figure for Poland was about 57%. The figures for the other countries lie somewhere between these two extremes. Thus, the crisis vulnerability in the glass/ceramics/cement sector reflects to a large extent the overall vulnerability of economy at country level. However, the majority of firms in these sectors, about 50% (ceramics) to 59% (glass) of them, was not affected by the crisis with respect to ICT investment, thus showing a considerable crisis resistance. This could be interpreted as a hint that ICT-related investment behaviour in these sectors shows no cyclicity or even is anti-cyclical (see the discussion in section 2). We concentrate here to those firms (40% to 50% of all firms) that were affected by the crisis and showed a pro-cyclical behaviour as to ICT investment.

| | No | Yes, no cancel- ling / downsiz- ing ICT pro- jects | Yes, cancel- ling / down- sizing ICT projects | N total |
|----------------|------|---|--|------------|
| <i>Sector</i> | | | | |
| Glass | 59.1 | 15.1 | 25.8 | 159 |
| Ceramics | 51.6 | 23.2 | 25.2 | 155 |
| Cement | 58.5 | 20.2 | 21.3 | 362 |
| <i>Country</i> | | | | |
| Germany | 74.4 | 8.9 | 16.7 | 180 |

| | | | | |
|--|------|------|------|-----|
| Spain | 44.0 | 25.6 | 30.4 | 125 |
| France | 65.0 | 14.0 | 21.0 | 86 |
| Italy | 47.5 | 21.8 | 30.7 | 101 |
| United Kingdom | 64.0 | 21.5 | 14.1 | 64 |
| Poland | 43.4 | 30.8 | 25.8 | 120 |
| <i>Firm size (number of employees)</i> | | | | |
| 10 – 49 employees | 65.1 | 19.2 | 15.7 | 364 |
| 50 – 249 employees | 51.9 | 19.3 | 28.8 | 264 |
| 250 employees and more | 36.1 | 26.7 | 37.2 | 86 |

Table 1. Impact of economic crisis on ICT investment plans or ICT projects

4 Methodology - Models Specification

In order to test our research hypotheses (section 2.3) we estimated some models, whose specification is described in this section, together with some econometric issues concerning their estimation. The exact definitions of all dependent and independent variables (= corresponding questions of the e-Business Survey 2009) are shown in the Appendix.

4.1 Explanation of the extent of ICT-related crisis vulnerability (crisis equation)

Dependent variable. We construct a binary variable (ICT_CRISIS) based on the following question of the e-Business Survey 2009: “Has the economic crisis an impact on your ICT investment plans or on ICT projects?”. The variable is codified as follows: 0: no impact; 1: yes, no ICT or e-business projects “were cancelled or significantly downsized” or yes, “ICT or e-business projects were cancelled or significantly downsized”.³

Independent variables. We distinguish three groups of variables:

ICT inputs: We use four variables, two related to the availability of ICT-specific human resources (ICT_personnel; ICT_skill deficits; both of them binary variables), one concerning the available financial resources (ICT_budget; a three-level ordinal variable), and one referring to the existence of ICT-outsourcing, meant as use of external ICT services as a measure to reduce ICT-related costs.

ICT infrastructure: The use of e-commerce is one important ICT application that serves to manage the firm functions as consumer (e-procurement E_P) as well as supplier of goods and services (e-sales E_S). Both e-commerce variables are six-level ordinal variables. In the Survey information is also available for a wide spectrum of other standard ICT functions and applications, such as Enterprise Resource Planning (ERP), Supply Chain Management (SCM), Supplier Relationship Management (SRM) as well as Customer Relationship Management (CRM). Further applications (CAD, CAM) refer to manufacturing activities in the more narrow sense. We construct separate variables for firms that use specific applications by way of ‘software-as-a-service’ (SAAS) and firms for which one or more

³ We constructed also a three-level ordinal variable showed similar results as the binary one. We prefer to show here the estimates for the simpler binary variable.

of the following new technologies are of relevance (in the sense that firms are aware of and have interest in them): service-oriented architectures, web 2.0 applications, data warehouses, and mobile services such as mobile commerce and remote access technologies (ICT_NEWTECH). Firms' ICT infrastructure is composed of all these technologies; the respective binary variables are included in our model.

Market conditions: We distinguish three dimensions of competition at the product market: price competition (PCOMP), quality competition (QCOMP) and customer service competition (SCOMP). The respective measures of the importance of these types of competition for the firms are three-level ordinal variables. Firms with intensive price competition in their main markets are likely to have difficulties to finance their ICT activities, since their price-cost margins are expected to be low. In good business times they are expected to have less problems to finance ICT investment. Hence, their investment behaviour is expected to be pro-cyclical, in our case meaning a positive correlation with the crisis variable. We expect that the competition pressure with respect to quality and customer service would be less strong than price competition in a crisis period. In this sense, firms that are exposed stronger to non-price competition are expected to be less crisis-vulnerable.

Other firm characteristics: Controls are inserted in the estimation equation for firm size, subsector, country, strong export-orientation (EXPORT) and the firm being part of a multinational corporation (INTER). At first glance, we would expect larger firms and firms belonging to multinational corporations to be less crisis-vulnerable than smaller firms because of the availability of more financial resources that help them to survive through a crisis. The descriptive data in table 1 show that it is the other way around, small firms being less crisis vulnerable than middle-sized and large firms, presumably because (i) larger firm are stronger dependent than small firms on exports that have decreased significantly as a consequence of the crisis, and (ii) small firms make less ICT investments, usually limited to the absolutely necessary ones, so there are less opportunities for ICT investment due to the crisis.

The economic crisis that started in 2008 has led to a decrease of world trade volume, thus to exports. Export-oriented firms have been stronger exposed to this particular crisis component than firms that primarily cover the domestic demand. Thus, we expect that export-oriented firms of the glass, ceramics and cement sector, which is primarily oriented to domestic demand, would be particularly affected by the crisis.

As already mentioned, the analysis of the ICT characteristics that may influence the likelihood of having reduced ICT investment as a consequence of the crisis is primarily explorative, thus driven by available data and economic intuition. A formal expression of the crisis equation is as follows:

$$ICT_CRISIS_i = a_0 + a_1ICT_budget_i + a_2ICT_outsourcing_i + a_3ICT_personnel_i + a_4ICT_skills_deficit_i + a_5E_Pi + a_6E_Si + a_7ERP_i + a_8SCM_i + a_9CRM_i + a_{10}CAD + a_{11}CRM + a_{12}SAAS + a_{13}ICT_NEWTECH + a_{14}PCMP + a_{15}QCOMP + a_{16}SCOMP_i + a_{17}INTER_i + a_{18}EXPORT_i + a_{19}Medium-sized_i + a_{20}Large_i + sector\ controls + country\ controls + e_i \quad (1) \text{ (for firm } i)$$

4.2 Effect of ICT-related crisis vulnerability and ICT characteristics on ICT-enabled innovation performance (innovation equations)

Dependent variables. We exploit the available information in the survey concerning the introduction of product and process innovations in the 12 months before the survey that are closely related to ICT and distinguish six different measures of ICT-related product and process innovation: (a) the new products or services contain *ICT components*; (b) new products or services for which ICT has played an important part in the *R&D process*, which led to the new products or services; (c) new products or services for which ICT has played an important part in the *market launch* of the new products or services; (d) new processes that are supported by ICT; (e) new processes for which ICT has played an important part in *the process design*; (f) new processes for which ICT has played an important part in

the implementation of the new processes. These variables allow a differentiated measurement of the specific contribution of ICT to product and process innovation.

Independent variables. We use the same vector of variables as in the crisis equation, with the exception of the variable *ICT_skills_deficits* as in the crisis equation, in addition we include the variable *ICT_CRISIS*. We expect a negative correlation of the crisis variable on all six measures for *ICT-enabled innovation* (*hypothesis 2*). According to *hypothesis 3* we expect a *joint positive* effect of the three groups of variables representing ICT inputs and ICT infrastructure, respectively, but we do not dispose of more detailed hypotheses regarding the effects of each single variable in the abovementioned groups of ICT variables. Finally, we expect a joint positive effect for three competition variables and a positive effect for firm size according to standard economic theory and empirical evidence (see, e.g., Cohen 2010 for a comprehensive survey of studies on innovation economics at the firm level). Formally expressed our innovation equations are:

$$[INNOPD_ICTx; INNOPC_ICTx]_i = a_0 + a_1ICT_CRISIS + a_2ICT_budget_i + a_3ICT_outsourcing_i + a_4ICT_personnel_i + a_5E_Pi + a_6E_Si + a_7ERP_i + a_8SCM_i + a_9CRM_i + a_{10}CAD + a_{11}CRM + a_{12}SAAS + a_{13}ICT_NEWTECH + a_{14}PCMP + a_{15}QCOMP + a_{16}SCOMP_i + a_{17}INTER_i + a_{18}EXPORT_i + a_{19}Medium-sized_i + a_{20}Large_i + sector\ controls + country\ controls + e_i (2) \text{ (for firm } i)$$

4.3 Econometric issues

Activities directed to product innovation and those aiming at process innovation are closely related (see, e.g., Athey and Schmutzler 1995 for a theoretical justification of this close complementary relationship; Kraft 1990; and Rouvinen 2002 for empirical evidence). In order to take this interdependence into account we estimated a trivariate model for *INNOPD_ICT1*, *INNOPD_ICT2* and *INNOPD_ICT3* as well as a trivariate model for *INNOPC_ICT1*, *INNOPC_ICT2* and *INNOPC_ICT3*. The correlation coefficients (values between 0.5 and 0.9) that were estimated for the trivariate models show that the choice of a trivariate model is econometrically justified. Due to technical difficulties we could not estimate a multivariate model for all six variables for ICT-enabled product and process innovation. This is not a big problem because the interdependence between the three ICT-enabled product innovation variables and between the three ICT-enabled process innovation variables is considerably larger than between product and process innovation as the correlation coefficients of additional estimations not shown here demonstrate.

We applied the STATA-procedure 'mvprobit', which estimates M-equation probit models, by the method of maximum simulated likelihood (MSL). The variance-covariance matrix of the cross-equation error terms has values of 1 on the leading diagonal, and the off-diagonal elements are correlations to be estimated ($\rho_{ji} = \rho_{ij}$, and $\rho_{ii} = 1$, for all $i = 1, \dots, M$). The procedure uses the Geweke-Hajivassiliou-Keane (GHK) simulator to evaluate the M-dimensional Normal integrals in the likelihood function. For each observation, a likelihood contribution is calculated for each replication, and the simulated likelihood contribution is the average of the values derived from all the replications. The simulated likelihood function for the sample as a whole is then maximized using standard methods (maximum likelihood in this case). For a brief description of the GHK smooth recursive simulator, see Greene (2011) who also provides references to the literature.

For the binary variable *CRISIS* we used a maximum likelihood probit model (STATA 'probit').

Due to the cross-section character of our data, both the left-hand and the right-hand variables refer to the same time period. As a consequence, our estimates of both the innovation and the performance equations have to be seen primarily as an extensive analysis of the correlations between the determinants (that are considered as structural characteristics that change only slowly over time) and the innovation and performance indicators, respectively. Nevertheless, some robust regularities come out, which if interpreted in view of our hypotheses presented in section 2 could possibly indicate the direction of causal links.

However, the possible distorting effect of the crisis variable, which is explicitly considered to be an endogenous variable in equation (1), in the innovation equations summarized in (2) has to be taken into account. To this end, we applied the procedure developed by Rivers and Vuong (1988) to test the exogeneity of the crisis variable in the innovation equations. We calculated the residual (predicted value minus effective value of variable ICT_CRISIS) of equation (1) and inserted it as a further right-hand variable in the innovation equations summarized in (2). As instrument we used the variable ICT_skill_deficits, which is significantly correlated with CRISIS but not with any of the ICT-enabled innovation variables and is not included in the specification of (2). Lack of ICT skills could at least partially be a problem due to limited supply of ICT-skilled personnel in some countries, thus exogenous to firms. The coefficients of the residual variable were not statistically significant at the 10% test level in any of the innovation equations. As a consequence, the corresponding estimates are based on the original crisis variable. The detailed results of the exogeneity tests are available upon request.

Multicollinearity is not an issue in our estimations as shown in the correlation matrix we estimated. Marginal effects were not estimated because 14 out of 21 right-hand variables in (1) and (2) are dummy variables and the rest are 7 ordinal variables.

5 Results

5.1 Crisis equation

In Table 2 we can see the estimated ICT-related crisis vulnerability model. We remark that it correlates positively with decreasing ICT budgets (pro-cyclical investment behaviour), the existence of skill deficits in ICT (as they reduce the value that can be generated for the firm from new ICT investments), the degree of perceived relevance of novel ICT applications (ICT_NEWTECH) that presumably request much additional ICT investment, the exposure to strong price competition (PCOMP) and the strong presence in international markets (EXPORT), in which activities have significantly decreased due to the crisis. None of the coefficients of the variables for the standard ICT applications is statistically significant, thus indicating that firms' ICT technology portfolio is not related to crisis vulnerability. These findings seem to support partially hypothesis 1b, with respect to ICT inputs (as lower levels of ICT budget and skill deficits increase ICT-related crisis vulnerability), and partially hypothesis 1c with respect to ICT infrastructure.

The positive sign for ICT_NEWTECH can be interpreted as a hint that firms that are more aware of the relevance of new technologies, and as a consequence of the need for much additional investment for exploiting them, tend to cancel or significantly downsize planned ICT projects (that concern mainly 'traditional' technologies) in order to fund the introduction in the future of new technologies, thinking that they have to exploit them to improve their ICT infrastructure as a reaction to crisis.

| | ICT_CRISIS |
|--------------------|----------------------|
| <i>ICT inputs</i> | |
| ICT_budget | -0.718*** (0.118) |
| ICT_outsourcing | 0.084 (0.149) |
| ICT_personnel | 0.149 (0.136) |
| ICT_skill deficits | 0.314*** |

| | |
|-----------------------------------|----------------------------|
| | (0.119) |
| <i>ICTinfrastructure</i> | |
| E_P | 0.042 (0.054) |
| E_S | -0.022 (0.060) |
| ERP | 0.081 (0.143) |
| SCM | 0.136 (0.185) |
| SRM | 0.122 (0.201) |
| CRM | -0.100 (0.152) |
| CAD | 0.103 (0.121) |
| CAM | -0.051 (0.153) |
| SAAS | 0.069 (0.151) |
| ICT_NEWTECH | 0.169*** (0.041) |
| <i>Market conditions</i> | |
| PCOMP | 0.295*** (0.101) |
| QCOMP | -0.169 (0.117) |
| SCOMP | 0.090 (0.091) |
| <i>Other firm characteristics</i> | |
| INTER | 0.205 (0.174) |
| EXPORT | 0.244* (0.149) |
| Medium-sized | -0.036 (0.130) |
| Large | 0.078 |

| | |
|-----------------------|-----------|
| | (0.205) |
| Sector dummies (2) | Yes |
| Country dummies (5) | Yes |
| Const. | -1.114*** |
| | (0.372) |
| N | 676 |
| Wald chi ² | 173.7*** |
| Pseudo-R ² | 0.192 |

Table 2. Ordered probit estimates for ICT_CRISIS (Note: Heteroskedasticity-robust standard errors in brackets; ***, **, *, denote statistical significance at the 1%, 5% and 10% test level resp.; reference for 'medium-sized', 'large': 'small').

5.2 Innovations equations: ICT-enabled innovation performance

5.2.1 Product innovation

The relationship between ICT-enabled product innovation and crisis vulnerability is negative for all three categories of ICT enablement, but only for INNOPD_ICT1 (new products contain ICT components) is this relationship statistically significant (column 1 in Table 3). The explanation for this negative effect is straightforward: the stronger the crisis impact on ICT investment, the less ICT-enabled innovation can be generated, particularly in the case of products and services that directly contain ICT components. This finding is in accordance with hypothesis 2 (pro-cyclical behaviour with respect to ICT-enabled innovation). However, the statistically significant negative relationship is found only for new products or services that contain ICT components, and are therefore directly affected by decreasing product demand. No significant relationship could be found for the indirect influence of ICT on product innovation through the R&D or marketing channel.

| | INNOPD_ICT1 | INNOPD_ICT2 | INNOPD_ICT3 |
|---------------------------|-------------|-------------|-------------|
| ICT_CRISIS | -0.546*** | -0.017 | -0.064 |
| | (0.217) | (0.199) | (0.207) |
| <i>ICT inputs</i> | | | |
| ICT_budget | 0.052 | 0.007 | -0.014 |
| | (0.169) | (0.170) | (0.170) |
| ICT_outsourcing | 0.423** | -0.100 | 0.029 |
| | (0.218) | (0.213) | (0.199) |
| ICT_personnel | 0.657*** | 0.710*** | 0.534*** |
| | (0.218) | (0.210) | (0.209) |
| <i>ICT infrastructure</i> | | | |
| E_P | 0.200*** | 0.231*** | 0.029 |
| | (0.081) | (0.087) | (0.079) |
| E_S | 0.045 | -0.068 | -0.015 |

| | | | |
|-----------------------------------|-----------|-----------|-----------|
| | (0.089) | (0.096) | (0.083) |
| ERP | -0.116 | -0.319 | 0.085 |
| | (0.257) | (0.233) | (0.233) |
| SCM | 0.187 | 0.012 | -0.026 |
| | (0.313) | (0.297) | (0.264) |
| SRM | 0.210 | -0.187 | -0.085 |
| | (0.312) | (0.280) | (0.272) |
| CRM | 0.299 | 0.152 | 0.228 |
| | (0.234) | (0.226) | (0.214) |
| CAD | 0.506** | 0.421** | -0.117 |
| | (0.210) | (0.201) | (0.193) |
| CAM | 0.302 | -0.228 | -0.233 |
| | (0.241) | (0.250) | (0.228) |
| SAAS | -0.308 | 0.255 | 0.022 |
| | (0.251) | (0.246) | (0.225) |
| ICT_NEWTECH | 0.150** | 0.129** | 0.175*** |
| | (0.063) | (0.059) | (0.059) |
| <i>Market conditions</i> | | | |
| PCOMP | 0.281 | 0.119 | 0.162 |
| | (0.170) | (0.157) | (0.148) |
| QCOMP | -0.054 | 0.370* | 0.219 |
| | (0.200) | (0.219) | (0.180) |
| SCOMP | 0.366** | -0.207 | -0.052 |
| | (0.165) | (0.157) | (0.142) |
| <i>Other firm characteristics</i> | | | |
| INTER | -0.138 | 0.185 | -0.165 |
| | (0.251) | (0.262) | (0.243) |
| EXPORT | 0.009 | -0.012 | 0.275 |
| | (0.286) | (0.251) | (0.234) |
| Medium-sized | 0.396 | 0.348 | 0.140 |
| | (0.245) | (0.218) | (0.213) |
| Large | 0.148 | 0.438 | -0.488 |
| | (0.331) | (0.300) | (0.289) |
| Sector dummies (2) | Yes | Yes | Yes |
| Country dummies (5) | Yes | Yes | Yes |
| Const. | -3.457*** | -2.622*** | -1.365*** |
| | (0.635) | (0.658) | (0.463) |

| | |
|---------------------------------------|----------|
| N | 247 |
| Wald chi ² | 229.8*** |
| Rho21 | 0.672** |
| Rho31 | 0.510*** |
| Rho32 | 0.795*** |
| Wald chi ² test of rho = 0 | 100.0*** |

Table 3. *Trivariate probit estimates for ICT-enabled product innovation variables INNOPD_ICT1, INNOPD_ICT2 and INNOPD_ICT3 (Note: Heteroskedasticity-robust standard errors in brackets; ***, **, * denote statistical significance at the 1%, 5% and 10% test level resp.; reference for 'medium-sized', 'large': 'small').*

ICT budget change is not relevant for all three types of ICT-enabled product innovation, presumably because the main effect of financing ICT is captured by the crisis variable. ICT outsourcing is positively correlated for innovation activities for which ICT components are contained in the new products. A possible explanation for this effect could be that ICT components may be cheaper acquired from specialized firms, so that own production of such components can be outsourced. Having specialized ICT personnel is significantly positively correlated with all three categories of ICT-enabled product innovation. Also, for all these three types of innovation is the degree of perceived relevance of novel ICT technologies (ICT_NEWTECH) positively correlated with innovation; this indicates that firms generating ICT-enabled product innovations have higher awareness of and show higher interest in novel ICT, in order to exploit them for further product innovations. With respect to standard ICT applications, e-procurement and CAD seem to be of relevance for the first two categories of product innovation (ICT components in products; ICT in R&D processes). Standard ICT applications are of no importance for ICT-enabled innovation that is related to the market launching of new products (column 3 in Table 3). On the whole, the variables for ICT inputs as well as ICT infrastructure show jointly positive effects on ICT-related product innovation, thus support hypothesis 3.

Price competition is not relevant for ICT-enabled product innovation. The two relevant dimensions of competition are quality competition (for INNOD_ICT1) and customer service competition (for INNOPD_ICT2), respectively. This finding is in accordance with our theoretical expectation. Belonging to a group of firms and being active at international markets are not important for ICT-enabled product innovation. A further interesting result refers to the absence of a size effect (with the exception of a positive effect for medium-sized firms in the estimates for INNOPD_ICT2): the likelihood of innovation in the glass/ceramics/cement industry appears to be independent of firm size, innovation activities are not concentrated in large firms, as it is often the case in other industries.

5.2.2 Process innovation

The relationship between ICT-enabled product innovation and crisis vulnerability is not statistically significant for all three categories of ICT enablement in process innovation, as we can see from Table 4. The decrease of ICT investment is not correlated with the introduction of ICT-enabled process innovation. This is not in accordance with the hypothesis 2 for pro-cyclical behaviour with respect to process innovation. Presumably, firms continued to improve their production processes in the crisis period using ICT resources in innovative ways more efficiently than before.

| | INNOPC ICT1 | INNOPC ICT2 | INNOPC ICT3 |
|---------------------------|--------------------|---------------------|---------------------|
| ICT_CRISIS | 0.128 (0.194) | -0.099 (0.197) | -0.126 (0.187) |
| <i>ICT inputs</i> | | | |
| ICT_budget | 0.021 (0.158) | 0.035 (0.166) | 0.170 (0.171) |
| ICT_outsourcing | 0.387** (0.196) | 0.757*** (0.201) | 0.827*** (0.219) |
| ICT_personnel | 0.403** (0.191) | 0.650*** (0.185) | 0.560*** (0.189) |
| <i>ICT infrastructure</i> | | | |
| E_P | 0.065 (0.084) | 0.091 (0.083) | 0.060 (0.084) |
| E_S | 0.100 (0.095) | 0.077 (0.086) | 0.048 (0.093) |
| ERP | 0.364 (0.228) | 0.169 (0.237) | 0.164 (0.236) |
| SCM | -0.058 (0.321) | -0.059 (0.334) | 0.124 (0.345) |
| SRM | 0.602* (0.336) | 0.073 (0.320) | 0.232 (0.343) |
| CRM | 0.112 (0.220) | -0.188 (0.213) | 0.122 (0.228) |
| CAD | 0.380** (0.190) | 0.066 (0.196) | 0.030 (0.188) |
| CAM | 0.098 (0.242) | -0.005 (0.229) | -0.030 (0.221) |
| SAAS | -0.059 (0.251) | -0.213 (0.250) | -0.383 (0.258) |
| ICT_NEWTECH | 0.134** (0.059) | 0.317*** (0.069) | 0.332*** (0.064) |
| <i>Market conditions</i> | | | |
| PCOMP | -0.091 (0.151) | -0.001 (0.147) | -0.086 (0.139) |
| QCOMP | -0.054 (0.210) | -0.094 (0.200) | 0.052 (0.173) |
| SCOMP | -0.196 | -0.130 | -0.216 |

| | | | |
|-----------------------------------|---------------------|---------------------|---------------------|
| | (0.135) | (0.161) | (0.147) |
| <i>Other firm characteristics</i> | | | |
| INTER | 0.431 (0.291) | 0.316 (0.250) | 0.212 (0.247) |
| EXPORT | -0.076 (0.225) | -0.168 (0.228) | -0.111 (0.233) |
| Medium-sized | 0.656*** (0.196) | 0.528*** (0.205) | 0.508*** (0.204) |
| Large | 0.339 (0.287) | 0.659** (0.306) | 0.642** (0.302) |
| Sector dummies (2) | Yes | Yes | Yes |
| Country dummies (4) | Yes | Yes | Yes |
| Const. | -0.552 (0.565) | -1.040* (0.566) | -1.073** (0.537) |
| N | | 291 | |
| Wald chi2 | | 295.7*** | |
| Rho21 | | 0.763*** | |
| Rho31 | | 0.831*** | |
| Rho32 | | 0.912*** | |
| Wald chi2 test of rho = 0 | | 180.3*** | |

Table 4. Trivariate probit estimates for ICT-enabled process innovation variables INNOPC_ ICT1, INNOPC_ ICT2 and INNOPC_ ICT3 (Note: Heteroskedasticity-robust standard errors in brackets; ***, **, * denote statistical significance at the 1, 5 and 10% test level resp.; reference for 'medium-sized', 'large': 'small').

There is a common pattern for all three categories of ICT-enabled process innovation: availability of ICT-specialized personnel and ICT outsourcing play an important role for ICT-enabled process innovation. Presumably, the ICT needed for process innovation is less firm-specific than for product innovation, so that outsourcing becomes a valid alternative to the development of firm-specific ICT-applications inside the firm. The perceived interest in novel technologies is also positively associated with ICT-enabled process innovation; this indicates that firms introducing ICT-enabled process innovations have higher awareness of and show higher interest in novel ICT, in order to exploit them for further process innovations. For new process that are supported in general by ICT (INNOPC_ ICT1) standard applications such as "Supplier Relationship Management" (SRM) and CAD are also needed. On the whole, the variables for ICT inputs as well as ICT infrastructure show jointly positive effects also on ICT-related process innovation, providing support for hypothesis 3.

Competition does not seem to be of any relevance for all three types of ICT-enabled process innovation, which is not what we would theoretically expect especially with respect to price competition. Larger firms show a significantly stronger tendency to ICT-enabled process innovation than medium-sized firms, and medium-sized firms show a stronger tendency than small firms. Thus, for process innovation, a size effect is found; bigger firm size leads to more complex processes, with numerous steps and involved employees, so there is stronger motivation for and benefit from exploiting the capabilities offered by ICT in order to generate process innovations.

6 Conclusions

6.1 Lessons Learned

From our empirical study the following lessons have been learned concerning our three research questions stated in the Introduction (at least with respect to the three manufacturing sectors (glass, ceramics and cement) from which our firm level data have been collected):

- *Crisis vulnerability*: The ICT-related crisis vulnerability correlates positively with decreasing ICT budgets (pro-cyclical investment behaviour), the existence of skill deficits in ICT, the perceived relevance of the potential use of new ICT applications that presumably request much additional ICT investment, the exposure to strong price competition and the strong presence in international markets, in which activities have significantly decreased due to the crisis,
- while on the contrary, none of the coefficients of the variables for the standard ICT applications is statistically significant, thus indicating that firms' ICT technology portfolio is not related to crisis vulnerability.
- *Crisis and innovation performance*: Statistically significant negative relationship between ICT-enabled product innovation and crisis vulnerability is found only for new products or services that contain ICT components; so only for this component of ICT-related innovation appears pro-cyclical behaviour to be dominant,
- while on the contrary no significant relationship could be found for the indirect influence of ICT on product innovation through the R&D or marketing channel.
- The decrease of ICT investment due to the crisis is not correlated with the introduction of ICT-enabled process innovation. Presumably, firms continued to improve their production processes in the crisis period using ICT resources more efficiently than before.
- *ICT characteristics and innovation performance*: Human resources in the form of ICT-specialized personnel are important for all six examined categories of ICT-enabled innovation.
- For four out of these six innovation categories ICT outsourcing is also relevant.
- The ICT technology portfolio characterized by a series of ICT applications is only partially important for innovation performance; the use of single applications such as E_P and CAD correlate positively with two out of the three examined product innovation categories, while CAD and SRM are important for one process innovation category.
- A common innovation-relevant characteristic for all six examined ICT-enabled innovation categories is the awareness of and interest in novel technologies, such as service-oriented architectures, web 2.0 applications, data warehouse and data mining, and mobile services such as mobile commerce and remote access technologies.

6.2 Implications - Future Research Directions

The discussion about the pro-cyclical or anti-cyclical character of different categories of investment that are associated with high risks and high sunk costs, e.g., R&D investment, or are of specific relevance for innovation output, e.g., ICT investment, is important not only for management but also for policy-making. For example, anti-cyclical investment behaviour with respect to ICT might mitigate negative implications of the cyclical movement of the economy both at firm level (e.g. high costs for search and recruitment of high-qualified ICT) and at economy level (e.g. deterioration of ICT infrastructure). Therefore the abovementioned findings – lessons learned from our study have interesting implications for practice. Also, they have implications for research, as they open up interesting research directions concerning the effect of economic crises (which are not rare in our economic system)

on firm's ICT investment and through it on innovation activity; our study provides a framework for this research.

However, further research is required in this direction, using data from several sectors of the economy, in order to be able to better understand the mechanisms behind cyclical behaviour of ICT investment. Furthermore, it would be interesting to examine the effects of crises on various types of 'hard' ICT investment (e.g. on various 'classical' and emerging types of ICT hardware and software) and 'human' ICT investment (e.g. on the employment and training of ICT personnel), and also on firms' ICT processes and structures (e.g. are there any changes of them in order to improve their efficiency and effectiveness?).

References

- Arvanitis, S. and M. Woerter (2014). "Firm Characteristics and the Cyclicity of R&D Investments". *Industrial and Corporate Change* 23 (5), 1141-1169.
- Barlevy, G. (2007). "On the Cyclicity of Research and Development". *American Economic Review* 97 (4), 1131-1164.
- Brynjolfsson, E. and A. Saunders (2010). *Wired for Innovation: How Information Technology Is Reshaping the Economy*. Cambridge, Mass: MIT Press.
- Cohen, W. M. (2010). "Fifty Years of Empirical Studies of Innovative Activity and Performance". In *Handbook of the Economics of Innovation - Volume 1*. Ed. by B.H. Hall and N. Rosenberg. Amsterdam: Elsevier B.V., pp. 129-213.
- Enkel, E., Gassmann, O. and H. Chesbrough (2009). "Open R&D and Open Innovation: Exploring the Phenomenon". *R&D Management* 39 (4), 311-316.
- Filippetti A. and D. Archibugi (2011). "Innovation in Times of Crisis: National Systems of Innovation Structure and Demand". *Research Policy* 40, 179-192.
- Gerner, H. D. and J. Stegmaier (2013). "Investitionen in der Krise? Eine empirische Analyse zum Einfluss der Finanz- und Wirtschaftskrise 2008/2009 auf Investitionsanpassungen". *Schmollers Jahrbuch*, 133, 67-96.
- Geroski, P.A. and P. Gregg (1997). *Coping with Recession – UK Company Performance in Adversity*. Cambridge, Mass: Cambridge University Press.
- Goodacre, A. and I. Tonks (1995). "Finance and Technological Change: In *Handbook of the Economics of Innovation and Technological Change*. Ed. by P. Stoneman. New York: Harvester.
- Greene, W.H. (2011). *Econometric Analysis - 7th ed*. New York: Prentice-Hall.
- Guellec, D. and E. Ioannidis(1999). "Causes of Fluctuations in R&D Expenditures: A Quantitative Analysis". *OECD Business Studies No. 29*, 1997/II.
- Kahle, K.M. and R. M. Stulz (2012). *Access to Capital, Investment, and the Financial Crisis*. Fisher College of Business Working Paper No. 2012-2, available at SSRN.
- Kleis, L., Chwelos, P., Ramirez, R. and I. Cockburn (2012). "Information Technology and Intangible Output: The Impact of IT investment on Innovation Productivity". *Information Systems Research* 23(1), 42-59.
- Leidner, D.E., Beatry, R.C. and J.M. Mackay (2003). "How CIOs Manage IT during Economic Decline: Surviving and Thriving amid Uncertainty". *MIS Quarterly Executive* 2(1), 1-14.
- OECD (2009). *The Impact of the Crisis on ICTs and Their Role in the Recovery*. Paris, OECD Publishing.
- OECD (2010): "Recent Developments and Outlook", in: *OECD Information Technology Outlook 2010*, OECD Publishing, Paris.
- Ouyang, M. (2011). "On the Cyclicity of R&D". *Review of Economics and Statistics* 93(2), 542-553.
- Rafferty, M. and M. Funk (2004) "The Effect of Demand Shocks on Firm-Financed R&D". *Research in Economics* 58, 187-203.
- Rivers, D. and Q.H. Vuong (1988). "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models". *Journal of Econometrics* 39, 347-366.

- Rojko, K., Lesjak, D. and V. Vehovar (2010). "Information Communication Technology Spending in (2008-) Economic Crisis". *Industrial Management & Data Systems* 111(3), 391-409.
- Sectoral e-Business Watch (2009). *ICT and e-Business Impact in the Glass, Ceramics and Cement Industry*. Study No. 02 conducted by empirica GmbH, Bonn and Brussels.
- Stiglitz, J.E. and A. Weiss (1981). "Credit Rationing in Markets with Imperfect Information". *American Economic Review* 71, 393-410.
- Tapscott, D., Ticoll, D. and A. Lowi (2000). *Digital Capital – Harnessing the Power of Business Webs*. Boston: Harvard Business School Press.

Appendix

Definitions of variables

| Variables | Definition |
|----------------------------|---|
| <i>Dependent variables</i> | |
| ICT_CRISIS | Has the economic crisis an impact on your ICT investment plans or on ICT projects? Binary variable: 1: yes, no ICT or e-business projects were cancelled or significantly downsized or yes, ICT or e-business projects were cancelled or significantly downsized; 0: No impact |
| INNOPD ICT1 | The new products or services have <i>ICT components</i> ; binary variable: 1: yes; 0: /no |
| INNOPD ICT2 | ICT has played an important part in the <i>R&D process</i> which led to the new products or services; binary variable: 1: yes; 0: no |
| INNOPD ICT3 | ICT has played an important part in the <i>market launch</i> of the new products or services; binary variable: 1: yes; 0: no |
| INNOPC ICT1 | The new processes are supported by ICT; binary variable: 1: yes; 0: no |
| INNOPC ICT2 | ICT has played an important part in the <i>process design</i> ; binary variable: 1: yes; 0: no |
| INNOPC ICT3 | ICT has played an important part in the <i>implementation</i> of the new processes; binary variable: 1: yes; 0: no |

| | |
|--|--|
| <p><i>Independent variables</i></p> <p>ICT_budget</p> <p>ICT_personnel</p> <p>ICT_skill deficits</p> <p>E_P</p> <p>E_S</p> <p>ERP</p> <p>SCM</p> <p>SRM</p> <p>CRM</p> <p>CAD</p> <p>CAM</p> <p>SAAS</p> <p>ICT_NEWTECH</p> <p>PCOMP</p> | <p>Change of ICT budget (covering primarily current variable costs and depreciation): 3-level ordinal variable: 0: decrease; 1: no change; 2: increase</p> <p>Employment of ICT practitioners; yes/no</p> <p>Have the employees problems because of insufficient computer skills? yes/no</p> <p>Use of <i>e-procurement</i>: 6-level ordinal variable: 0: no e-procurement; 1: less than 5% of purchases; 2: 5-10%; 3: 11-25%; 4: 26-50%; 5: 50% and more</p> <p>Use of <i>e-sales</i>: 6-level ordinal variable: 0: no e-sales; 1: less than 5% of sales; 2: 5-10%; 3: 11-25%; 4: 26-50%; 5: 50% and more</p> <p>Use of 'Enterprise Resource Planning'; yes/no</p> <p>Use of 'Supply Chain Management'; yes/no</p> <p>Use of 'Supplier Relationship Management'; yes/no</p> <p>Use of 'Customer Relationship Management'; yes/no</p> <p>'Computer Aided Design'; yes/no</p> <p>'Computer Aided Manufacturing'; yes/no</p> <p>Use of specific applications by way of 'software-as-a-service', which is licensed for use on line; yes/no</p> <p>Relevance for the company of: service-oriented architectures; web 2.0 applications; data warehouses and data mining; mobile services such as mobile commerce and remote access technologies</p> <p>Importance of <i>price competition</i> in the main market; 3-level ordinal variable: 0: not important; 1: quite important; 2: very important</p> |
|--|--|

| | |
|-----------------|---|
| QCOMP | Importance of <i>quality competition</i> in the main market; 3-level ordinal variable: 0: not important; 1: quite important; 2: very important |
| SCOMP | Importance of <i>customer service competition</i> in the main market; 3-level ordinal variable: 0: not important; 1: quite important; 2: very important |
| INTER | Part of a <i>multinational</i> enterprise; yes/no |
| EXPORT | <i>International</i> market as most important sales market; yes/no |
| Medium-sized | 50 to 249 employees |
| Large | 250 employees and more |
| Sector dummies | Ceramics, cement (reference: glass) |
| Country dummies | France, Italy, Poland, Spain, United Kingdom (reference: Germany) |

LOUD COMPUTING ADOPTION MOTIVATION IN THE EUROPEAN NORTH AND SOUTH

Completed Research

Loukis, Euripidis, University of Aegean, Samos, Greece, eloukis@aegean.gr

Kyriakou, Niki, University of Aegean, Samos, Greece, nkyr@aegean.gr

Abstract

The European North-South divide has been one of the most important and widely debated problems of Europe for long time. The countries of the European South have for decades lower levels of economic and technological development and performance than the countries of the European North; though there has been a convergence between the European North and South for some time, recently, due to the economic crisis, this trend has stopped, and on the contrary a divergence is observed. It is widely recognized that in order to overcome this negative situation, and achieve a gradual convergence between these two regions, it is important not only to cut wages and public expenditure in the European South (which has been the dominant approach so far), but also to make wider and better use of new technologies and boost innovation. This paper contributes in this direction, comparing empirically the European North and South with respect to one of the most important, innovative and disruptive new information and communication technologies (ICT): the cloud computing (CC). CC is emerging as a new paradigm of providing ICT support of firms' and activities, which can not only reduce costs (especially investments), but also enable the rapid and low cost experimentation with and exploitation of new emerging technologies, and also support and facilitate innovation and external collaboration. In particular, we empirically investigate and compare Northern and Southern European firms with respect not to the 'quantity' of CC use, but to its 'quality': their CC adoption motivations and orientations. Our study has been based on a dataset collected through the e-Business W@tch Survey of the European Commission from 556 European firms from the glass, ceramic and cement sectors. It has been concluded that Southern European firms of the above sectors view CC as a means of reducing ICT investment, while Northern European ones view it as a means of supporting and facilitating product/service innovation, and also of reducing cost and increasing capabilities of their existing external electronic collaboration (with business partners and experts) for the development of innovations. Furthermore, both Northern and Southern European firms of the above sectors view CC as a means of experimentation with and exploitation of new emerging technologies.

Keywords: cloud computing, adoption, innovation, European North-South divide.

1 Introduction

The European North-South divide has been one of the most important and widely debated problems of Europe for long time (Aiginger, 2013a and 2013b; Landesmann, 2013). The countries of the European South (often referred to as the ‘European Periphery’) have for decades lower levels of economic and technological development, productivity and performance, and also higher levels of unemployment, than the countries of the European North. The Southern European countries are characterised by some fundamental weaknesses associated with the size and structure of manufacturing, deficits in innovation and education, deficits with respect to the exploitation of economy globalisation and the restructuring of the public sector. They have a larger share of low skill industries and a smaller share of higher skill ones; the technology driven industries are much smaller in comparison with the Northern European countries, and also declining. European periphery countries did not use the advantage of globalization despite being located by the sea and despite a history of global trade connections. Though there has been a convergence between the European North and South for some time, recently, due to the economic crisis, this trend has stopped, and on the contrary a divergence is observed (Aiginger, 2013a). It is widely recognized that in order to overcome this negative situation, and achieve a gradual convergence between these two regions, it is important not only to cut wages and public expenditure in the European South (which has been the dominant approach so far), but also to make wider and better use of new technologies and boost innovation, aiming at increase of productivity and growth.

Cloud Computing (CC) is one of the most important, innovative and disruptive new information and communication technologies (ICT), which changes radically the way firms access and use ICT for supporting their activities: part of the ICT support services required by firms are delivered not by their internal ICT units, but by external providers on an on-demand basis over the Internet, and users pay for the service as an operating expense, based on the real use of it, without having to make significant initial hardware and software investments, and also without having to incur operation, support and maintenance costs (Armbrust et al., 2010; Marston et al., 2011; Venders and Whitley, 2012). CC is defined by the US National Institute for Standards and Technology (NIST) as “[...] a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of computing resources (e.g. networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service provider interaction” (NIST, 2009). CC can provide significant benefits to firms: initially the ICT cost reduction was regarded as the most significant of them, and especially the reduction of the required ICT investments, by converting related capital investments (cap-ex) to operating costs (op-ex); however, it was soon realised that CC could provide, beyond these ‘first-level’ cost reduction oriented benefits, some additional ‘second-level’ significant transformation oriented benefits: it can enable the rapid and low cost experimentation with and exploitation of new emerging technologies, and also support and facilitate innovation and external collaboration (Etro, 2009; Brynjolfsson et al., 2010; Marston et al., 2011; Venders and Whitley, 2012). According to Armbrust et al. (2010) CC enables the quick implementation of new ICT-based ideas, as ‘developers with innovative ideas for new Internet services no longer require the large capital outlays in hardware to deploy their service or the human expense to operate it’.

In this paper we empirically investigate and compare Northern and Southern European firms with respect not to the ‘quantity’ of CC use, but to its ‘quality’: their CC adoption motivations and orientations. In particular, we investigate and compare to what extent Northern and Southern European firms view CC as a means of: a) ICT investment reduction; b) Supporting and facilitating product/service innovation and process innovation; c) Experimenting with and exploiting new ICT; and d) Supporting and facilitating external collaboration. Furthermore, this investigation is not based on the analysis of firms’ managers’ subjective perceptions concerning the usefulness of CC along the abovementioned four dimensions (which is the usual approach in the research literature for investigating firms’ motivations for adopting various technologies); it adopts a more ‘objective’ approach, based on the estimation of the firm level association between propensity for CC adoption and each of the above four vari-

ables (adoption of ICT investment strategy, product/service innovation strategy and process innovation strategy; interest in some new emerging ICT (data warehousing and data mining, mobile services); and having external collaborations for the development of innovations) in these two geographic regions. Our study has been based on a dataset collected through the e-Business W@tch Survey of the European Commission from 556 European firms from the glass, ceramic and cement sectors.

This paper consists of seven sections. In the following section 2 a relevant literature review is presented, while in section 3 our research hypotheses are developed. Then in section 4 the data and the method of this study are described. In section 5 the results are presented and discussed. The final section 6 summarizes the conclusions and suggests future research directions.

2 Literature Review

Considerable empirical research has been conducted concerning the factors affecting CC adoption by firms. Most of it has been based on the Technology, Organization and Environment (TOE) theory (Tornatzky and Fleischer, 1990; Baker, 2011), which identifies three groups of factors that affect the adoption of technological innovations in general by firms: technological (= perceived characteristics of the technological innovation), organizational (= firm's characteristics) and environmental (= characteristics of firm's external environment) ones (Low et al., 2011; Mangula et al., 2014; Hsu et al., 2014; Oliveira et al., 2014; Gutierrez et al., 2015); for each group a number of CC related factors are determined, and then the statistical importance of their effects on adoption CC are tested. In particular, Low et al. (2011) examine the effect of a set of technological factors (relative advantage, complexity and compatibility), organizational factors (top management support, firm size and technology readiness) and environmental factors (competitive pressure and trading partner pressure) on CC adoption. They concluded that perceived relative advantage, top management support, firm size, competitive pressure and trading partner pressure have statistically significant effects on CC adoption. Mangula et al. (2014) similarly examine the effect of technological factors (relative advantage, compatibility, complexity, trialability, observability), organizational factors (organizational readiness, top management support) and environmental factors (market pressure, market competition vendor marketing, trust in vendor, government support) on the adoption of Software as a Service (SaaS) services. They conclude that compatibility, observability, market competition and government support have a positive correlation with SaaS adoption, while complexity has a negative correlation with it. Hsu et al. (2014) examine the effect of perceived benefits and business concerns (technological factors), IT capability (IT personnel and budget - organizational factor) and external pressure (environmental factor) on CC adoption intention; they found that the first three of these factors are significant determinants of CC adoption. Oliveira et al. (2014) examine the effects of three CC characteristics from a technological innovation perspective (relative advantage, complexity and compatibility), three organizational context characteristics (top management support, firm size, technological readiness) and two environmental context characteristics (competitive pressure, regulatory support). They found that relative advantage, technological readiness, top management support and firm size have positive effects on CC adoption, while complexity has a negative effect. Another similar study has been conducted by Gutierrez et al. (2015), who examined the effects of a set of technological factors (relative advantage, complexity and compatibility), organizational factors (top management support, firm size, technological readiness) and environmental factors (competitive pressure, trading partners pressure) on CC adoption. They concluded that competitive pressure, complexity, technology readiness and trading partner pressure have a significant influence on the adoption of CC services.

Furthermore, there are CC adoption empirical studies that are based on the synthesis of other theoretical frameworks (Benlian et al., 2009; Saya et al., 2010; Benlian & Hess, 2011; Wu et al., 2013). The study of Benlian et al. (2009) developed a SaaS adoption model by combining three theoretical foundations: transaction cost theory (including in their model the application specificity and perceived uncertainty), resource-based view of the firm (including application strategic value and inimitability) and

theory of planned behaviour (including the attitude towards SaaS and also social influence). It has concluded that social influence, adoption uncertainty and application strategic value are the most consistent SaaS adoption drivers across all application types. The study of Saya et al. (2010) formulated a four layers structural equation CC adoption model, based on the institutional theory and the real options theory. It reached the conclusion that institutional influences (e.g. from government, customers, suppliers, competitors, strategic partners, industry and trade organizations, professional bodies) affect organizations perceptions about the technological characteristics of CC (perceived accessibility, scalability, cost effectiveness and lack of security), and through them affect the perceptions on the provided real options by CC adoption (concerning ICT applications growth, abandonment and deferral) and finally the intention to adopt CC). Benlian and Hess (2009) having as theoretical foundation the theory of reasoned action, in combination with previous research on ICT outsourcing and application service provision (ASP), examine the effects of perceived SaaS opportunities (cost advantages, strategic flexibility, focus on core competencies, access to specialised resources and quality improvements) and SaaS risks (performance, economic, strategic, security and managerial ones) on the intention to increase the level of its adoption. They have concluded that the perceived cost advantages have the strongest positive effect, followed by strategic flexibility and the quality improvement, while the security risks have the strongest negative effect, followed by the performance, economic and strategic risks. Wu et al. (2013) conducted a study of CC adoption factors having as theoretical foundations the innovation diffusion theory (DOI) proposed by Rogers (2003) and the 'information processing view' (IPV) of the firm. They concluded that business process complexity and also applications compatibility have negative effects on CC adoption intention, while entrepreneurial culture and applications functionality have positive effects.

However, there is a lack of empirical investigation of the association between CC adoption and various aspects of firm's strategy and operation, which would provide valuable insight into CC adoption motivations and orientations of firms. Also, there is a lack of comparative studies in this respect between geographic regions or countries, from such a 'quality' related perspective, rather than a 'quantity' related one. This study contributes to filling these research gaps, by empirically investigating and comparing Northern and Southern European glass, ceramic and cement sector firms with respect to their CC adoption motivations and orientations.

3 Research Hypotheses

We formulated four research hypotheses, which concern the association of the propensity to adopt CC with four firm's strategic and operational characteristics that correspond to the four main motivations of firms for adopting CC according to previous literature (Armbrust et al., 2010; Marston et al., 2011; Venders and Whitley, 2012): ICT investment reduction, support and facilitation of innovation, experimentation with and exploitation of new emerging ICT, and support and facilitation of external collaborations.

In particular, our first research hypothesis concerns the association between the adoption of an ICT investment reduction strategy and the propensity for CC adoption. Due to the economic crisis that exists in many countries firms have to adopt to a greater or lesser degree strategies of IT investment reduction. This does not allow them to upgrade and enhance their ICT infrastructures in order to meet new business needs, or to take advantage of new emerging technologies (such as data warehousing/mining, mobile technologies, etc.); this can have negative impact on firms' long term competitiveness. CC can be quite useful for such firms, as it enables them to upgrade the computing power of their ICT infrastructures (e.g. by using Infrastructure as a Service (IaaS)), and also their functionality (e.g. by using Software as a Service (SaaS)), without having to make additional upfront ICT investments (Marston et al., 2011; Venders and Whitley, 2012), transforming them to operational expenses based on the real use they make of these services (a 'pay as you go' model), and also without having to incur the corresponding operation, support and maintenance costs. Therefore, we expect that firms

adopting an ICT investment reduction strategy will have a strong propensity to adopt CC. So, our first research hypothesis is:

H1. *The adoption of an ICT investment reduction strategy is positively associated with the propensity for CC adoption.*

Our second research hypothesis concerns the association between the adoption of an innovation-oriented strategy and the propensity for CC adoption. Changes in customers' needs and preferences, emergence of new technologies and strong competition make it necessary for firms to make innovations in their products and services, and also in their internal production and administrative processes, which have become today highly important for the competitiveness and even for the survival of firms. However, these innovations (both product/service and process ones) usually necessitate the development of complex supporting ICT infrastructures; this can be costly (requiring considerable capital investments), risky (since if the innovation is not successful its supporting ICT infrastructure will become to a large extent useless, leading to waste of significant financial resources) and also can take too much time (which is quite negative in the rapidly changing and highly competitive modern economy). CC can alleviate the above problems: it can reduce the cost of the required ICT infrastructure for supporting an innovation (and make it an operational expense, without having to make ICT investments), and the implementation time (as the required CC services can be rapidly activated and customized), and also reduce the risk (since if the innovation is not successful the CC services used for supporting it can be simply terminated). Previous CC literature has emphasized that it can provide benefits associated not only with the ICT cost reduction, but also with the support and facilitation of innovations as well, as CC enables the rapid development of their required supporting ICT infrastructures, at a low cost, without requiring ICT capital investments (Brynjolfsson et al., 2010; Marston et al., 2011; Venders and Whitley, 2012; Berman et al., 2012). So, we expect that firms adopting an innovation-oriented strategy will have a strong propensity to adopt CC. Thus, our second hypothesis is:

H2. *The adoption of an innovation-oriented strategy is positively associated with the propensity for CC adoption.*

It can be analysed into the following two hypotheses:

H2a. *The adoption of a product or service innovation oriented strategy is positively associated with the propensity for CC adoption.*

H2b. *The adoption of a process innovation oriented strategy is positively associated with the propensity for CC adoption.*

Our third research hypothesis concerns the association between the interest in the adoption of new ICT and the propensity for CC adoption. A major trend of the modern economy is the continuous emergence of new ICT; each firm has to decide which of the multiple new emerging ICT are appropriate and beneficial for its particular activities, processes, products and services, and also sufficiently mature, so they should be adopted, and which of these emerging ICT are not, so they should not be adopted. However, the adoption of a new emerging ICT poses two important problems: on one hand it can be costly and require some capital investment, and on the other hand it carries some uncertainty and risk (as to whether it is really applicable, appropriate and beneficial); if it is not finally successful there will be a loss of valuable financial resources that have been used for the relevant investment. CC can alleviate both these problems: it can reduce the abovementioned required costs, making them operational expenses and eliminating the need for investment; also it can eliminate the inherent risk (since if the adoption is not successful the CC services used can be simply terminated). Previous literature argues that one of the most important advantages of CC is that it enables enhancing firm's ICT infrastructure by incorporating new emerging ICT, rapidly, at a low cost and without having to make additional investments, with the most widely mentioned of them being data warehousing/mining and mobile services (Marston et al., 2011; Venders and Whitley, 2012; Bhagyashree and Borkar, 2012; Verma, 2013). Therefore we expect that firms interested in experimentation with and exploitation of new ICT will have a strong propensity to adopt CC. So, our third hypothesis is:

H3. *Interest in adopting new ICT is positively associated with the propensity for CC adoption.*

Finally, our fourth research hypothesis concerns the association of the collaboration with other firms with the propensity to adopt CC. The globalization, the strong competition, the continuous emergence of new technologies, the fast changes that characterise the modern business environment, as well as the high expectations and demands of consumers for high value-added products and services, and also for continuous renewal and improvement of them, make it difficult for individual firms to survive on their own, relying only on their internal resources, and this results in increasing collaboration among firms having complementary resources, both at the operational and the product/service and process innovation level (Rycroft, 2007; Zeng et al., 2010; Xie et al., 2013; Majava et al., 2013). However, this necessitates extensive exchange of both structured and unstructured information, which can be significantly supported and facilitated through the use of appropriate ICT. The use of CC services enables the development, operation and maintenance of this ICT support of collaboration rapidly, at a low cost, without having to make additional investments. A recent study based on interviews with business and ICT practitioners in the UK revealed that CC has a strong potential to support and facilitate business collaboration at a low cost (Willcocks et al., 2014). For the above reasons we expect that firms having collaboration with other firms will have a strong propensity to adopt CC. So our fourth research hypothesis is:

H4. *Collaboration with other firms is positively associated with the propensity for CC adoption.*

4 Data and Method

In this study we have used firm level data collected through the “e-Business Survey 2009” survey, conducted by the e-Business Market W@tch (www.ebusiness-watch.org) under the auspices of the European Commission. In this survey were collected data concerning the use of various types of ICT, ICT skills, ICT investment, and also innovation activity from 676 firms of the glass, ceramic and cement sectors from six European countries: Germany, France, UK, Italy, Spain and Poland. For this study we focused on the data from the first five of these countries, with the first three belonging to the European North, and the next two belonging to the European South, while the Polish data were not used (as Poland belongs to the Easter European region, which has quite different economic characteristics).

As dependent variable we used the propensity for CC adoption, which has three possible values (very relevant, partly relevant, not relevant). As independent variables we used three binary variables concerning the adoption or not of an ICT investment reduction strategy, a product/service innovation strategy and a process innovation strategy respectively; two values concerning firm’s interest in two new emerging ICT, the data warehousing and data mining, and the mobile services, which have three possible values (very relevant, partly relevant, not relevant); and also two binary variables concerning the existence or not of external innovation collaboration (i.e. with other firms), and also of electronic (i.e. supported by ICT) external innovation collaboration.

We tested the research hypotheses H1 – H4 separately for the European North and the European South, by estimated the association between the dependent variable and each of the abovementioned independent ones, initially for the European North sub-sample (firms from Germany, France and Italy), and then for the European South sub-sample (Italy and Spain). In particular, we calculated widely used measures of association between ordinal variables: Sommer’s D and Kendall’s tau-b; they both range from -1 to 1, with the sign indicating the direction of the association, and the absolute value indicating its strength. It should be noted that we did not estimate a regression model because there were high levels of correlation between our independent variables; according to the econometric literature (e.g. Greene, 2011; Gujarati, 2008) if we have high levels of correlation between the independent variables of a regression (problem of ‘multi-collinearity’), then the regression coefficients are not reliable estimates of the impact of the independent variables on the dependent variable.

5 Results

In Table 1 we can see the relative frequencies of the three possible values of our dependent variable (propensity for CC adoption) for each of the five countries examined in this study. We remark that in the glass, ceramic and cement sectors of the examined Southern Europe countries there is a higher share of firms considering CC as very relevant or partly relevant than in the examined Northern European countries. A possible explanation of this might be that the economic problems of the European South limit the financial resources of firms, and this increases their propensity to use CC for reducing the ICT costs and especially ICT investments; however, a clearer picture on this can be formed by examining the estimated associations of the propensity for CC adoption with our independent variables, which are discussed in the following paragraphs of this section. In Table 2 we can see the relative frequencies of the possible values of our independent variables for each of the five countries examined in this study (for the two triple valued independent variables, concerning firm’s interest in data warehousing/data mining, and mobile services, is shown the sum of the relative frequencies of ‘very relevant’ and ‘partly relevant’). We remark in the European South, due to the existing economic problems that reduce demand and sales, it is much higher the percentage of firms of these sectors adopting an ICT investment reduction strategy, in comparison with the European North; also, we have higher percentages of firms making innovations, having electronic external innovation collaboration and being interested in data warehousing/mining.

| | | Cloud Computing Propensity | | |
|-------|---------|----------------------------|---------------------|------------------|
| | | Very Relevant (%) | Partly Relevant (%) | Not Relevant (%) |
| South | Italy | 4.0 | 17.8 | 8.2 |
| | Spain | 3.2 | 12.8 | 84.0 |
| North | UK | 0.0 | 4.7 | 95.3 |
| | France | 1.2 | 11.6 | 87.2 |
| | Germany | 0.0 | 4.4 | 95.6 |

Table 1. Cloud Computing propensity relative frequencies

| | | ICT Invest. Reduction Strategy | Product/Service Innovation Strategy | Process Innovation Strategy | Interest in DW/DM | Interest in Mobile Services | Innovation Collaboration | Electr. Innovation Collab. |
|-------|--------|--------------------------------|-------------------------------------|-----------------------------|-------------------|-----------------------------|--------------------------|----------------------------|
| South | Italy | 30.7 | 39.6 | 38.6 | 32.7 | 44.6 | 25.7 | 13.9 |
| | Spain | 30.4 | 36.0 | 44.0 | 26.4 | 57.6 | 13.6 | 13.6 |
| North | UK | 14.1 | 34.4 | 40.6 | 9.4 | 26.5 | 18.8 | 7.8 |
| | France | 20.9 | 20.9 | 24.4 | 22.1 | 41.9 | 17.0 | 11.6 |
| | Germ. | 16.7 | 36.1 | 39.4 | 17.2 | 33.9 | 16.1 | 9.4 |

Table 2. Independent variables frequencies

In Tables 3 and 4 we can see for all independent variables their calculated Sommer's D coefficient and Kendall tau-b coefficient values with respect to the independent variable (propensity for CC adoption) for the European North and South respectively (statistically significant values are shown in bold).

We remark that in both regions there is a statistically significant positive association of the interest in the two examined new emerging ICT, data warehousing/mining and mobile services, with the propensity for CC adoption. Therefore, hypothesis H3 is supported in both regions. This indicates that both Northern and Southern European firms of the examined sectors view CC as a low cost and risk means of experimentation with and exploitation of these new emerging ICT. We can see that in both regions the Sommer's D and Kendall tau-b values for these two variables are the highest among all our independent variables, which indicate that experimentation with/exploitation of emerging ICT seems to be a very strong motivation for the adoption of CC by firms of these three sectors in both Northern and Southern Europe; this appears to be stronger in the Southern Europe, as indicated by the higher values of both Sommer's D and Kendall tau-b for this region (see Table 4) in comparison with the Northern Europe (see Table 3), as the economic problems and the lower market demand in the former put pressure on them to exploit the extensive capabilities for low cost and risk use of new emerging ICT offered by CC.

| Independent Variable | Sommer's D | Kendall tau-b | Research Hypothesis |
|--|-------------|---------------|---------------------|
| ICT Investment Reduction Strategy | .094 | .102 | H1 |
| Product/Service Innovation Strategy | .044 | .053 | H2a |
| Process Innovation Strategy | .107 | .131 | H2b |
| Interest in Data Warehousing/Mining | .182 | .199 | H3 |
| Interest in Mobile Services | .124 | .155 | H3 |
| Innovation Collaboration | .063 | .067 | H4 |
| Electronic Innovation Collaboration | .276 | .280 | H4 |

Table 3. Sommer's D and Kendall tau-b values for the European North

| Independent Variable | Sommer's D | Kendall tau-b | Research Hypothesis |
|--|-------------|---------------|---------------------|
| ICT Investment Reduction Strategy | .205 | .208 | H1 |
| Product/Service Innovation Strategy | .113 | .116 | H2a |
| Process Innovation Strategy | .104 | .106 | H2b |
| Interest in Data Warehousing/Mining | .328 | .334 | H3 |
| Interest in Mobile Services | .232 | .245 | H3 |
| Innovation Collaboration | .118 | .118 | H4 |
| Electronic Innovation Collaboration | .112 | .113 | H4 |

Table 4. Sommer's D and Kendall tau-b values for the European South

Furthermore, we remark that only in the European South there is a statistically significant positive association of the adoption of an ICT investment reduction strategy with the propensity for CC adoption; so hypothesis H1 is supported only in the European South. On the contrary, only in the European North there is a statistically significant positive association of the adoption of a product/service inno-

vation strategy with the propensity for CC adoption; so hypothesis H2a is supported only in the European North. Also, there is not statistically significant association of the adoption of a process innovation strategy with the propensity for CC adoption; so hypothesis H2b is not supported. The above results indicate that the Southern European firms of the above sectors view CC as a means of reducing ICT investment; CC enables them to upgrade and enhance their ICT infrastructures in order to meet new business needs, without having to make new investments, which would be difficult to finance in the problematic economic context of the European South. On the contrary the Northern European firms of the above sectors view it as a means of supporting and facilitating product/service innovation.

Finally, in none of these two regions there is statistically significant positive association of the external innovation collaboration (i.e. collaboration with other firms for the development of innovations in products and services) with the propensity for CC adoption; only in the European North there is a statistically significant positive association of the electronic (i.e. supported by ICT) external innovation collaboration with the propensity for CC adoption. Therefore, hypothesis H4 is only partially supported. These results indicate that firms of these sectors from both regions do not view CC as a means of providing electronic support of innovation oriented collaboration with other firms, despite the arguments of relevant literature that CC has a strong potential to support and facilitate business collaboration at a low cost (Willcocks et al., 2014); however, in the European North firms already having electronically supported innovation collaboration with other firms view CC as a means of reducing the cost and increasing the capabilities and flexibility of the existing ICT support of their external innovation collaborations.

6 Conclusions

One of the most important problems of Europe for long time has been the gap in economic and technological development and performance between the European North and the European South, referred to as the 'European North-South divide'. Though for some time a gradual convergence between these two regions was in progress, recently, due to the economic crisis, this has stopped, and on the contrary a new divergence has started. It is widely recognized that in order to reverse this negative trend and achieve a gradual convergence between these two regions, it is of critical importance to make wider and better use of new technologies and boost innovation in the European South in order to improve its productivity. This study makes a contribution to this 'European North-South divide' debate, by empirically investigating and comparing European North and South with respect to the one of the most important, innovative and disruptive new ICT: the CC; this technology changes radically the way firms access and use ICT for supporting their activities, and also the economics of business computing (as it enables the conversion of relevant capital investments (cap-ex) to operating costs (op-ex). In particular, we investigate and compare the 'quality' (instead of the 'quantity' usually examined by similar studies) of CC use (or planned use) by the Northern and Southern European firms: we examine to what extent they view CC as a means of: a) ICT investment reduction; b) Supporting and facilitating product/service innovation and process innovation; c) Experimenting with and exploiting new ICT; and d) Supporting and facilitating external collaboration. Our study has been based on a dataset collected through the e-Business W@tch Survey of the European Commission from 676 European firms from the glass, ceramic, cement and sectors.

It has been concluded that in the European South firms of the above sectors have in general a higher interest in and propensity for the adoption of CC than in the European North. However, the motivations and orientations of the former with respect to CC adoption have some similarities with the ones of the latter, but also important differences as well. Both Northern and Southern European firms of the examined sectors view CC as a low cost and risk means of experimentation with and exploitation of new emerging ICT, and this seems for both regions to be the strongest CC adoption motivation among the examined ones. However, we have found for the Southern European firms a higher interest in the experimentation with/exploitation of the examined new emerging ICT in comparison with the North-

ern European ones (see Table 2), and also a stronger association of it with the propensity for CC adoption (see Tables 3 and 4), which indicates a stronger motivation/orientation of the Southern European firms to use CC for experimentation with/exploitation of the new emerging ICT. The economic problems and the lower market demand in the European South put pressure on firms to exploit the extensive capabilities for low cost and risk use of new emerging ICT offered by CC.

Furthermore, it has also been concluded that Southern European firms of the above sectors view CC as a means of reducing ICT investment; this does not hold for the Northern European ones, which on the contrary view CC as a means of supporting and facilitating product/service innovation, and also of reducing cost and increasing capabilities of their existing external electronic collaboration (with business partners and experts) for the development of innovations. This indicates that Southern European firms are mainly oriented towards 'first-level' cost (and especially investment) reduction related benefits from CC, while on the contrary Northern European firms are mainly oriented towards 'second-level' transformation related benefits from CC, which are associated with support and facilitation of innovation and external collaboration. The difficulty of financing investments in the problematic economic context of the European South, in combination with the longer and stronger tradition of the European North concerning the use and advanced exploitation of complex new technologies, are a possible explanation for this.

The findings of this empirical study have interesting implications both for research and practice. With respect to research it makes a contribution to the existing body of knowledge concerning the impact of the national context of ICT adoption, focusing on a very important and disruptive ICT (the CC), and also not on the 'quantity' of its adoption, but on its 'quality' (motivations and orientations of CC adoption). With respect to practice, our conclusions can be useful for government agencies, both at national level and at European level, in order to formulate effective technology adoption and transfer policies, and also for CC services providers, in order to optimize their offerings in taking into account the specific characteristics and needs of each national market. Our study has two main limitations: its limited sectoral and national scope, and also the use of a rather broad dependent variable (propensity for CC adoption in general). So further research is required concerning the motivations/orientations of the adoption of various types of CC services (e.g. IaaS, PaaS, SaaS), in various sectorial and national contexts.

References

- Aiginger, K. (2013a). "A New Strategy for the European Periphery". *Austrian Institute of Economic Research*. WIFO Working Papers - No. 443, Vienna, Austria.
- Aiginger, K. (2013b). "A Southern Europe strategy based on vision and industrial policy". *Ekonomiaz* 82 (1), 129-169.
- Armbrust, M., Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica I. and Zaharia M. (2010). "A view of Cloud Computing". *Communications of the ACM* 53 (4), 50-58.
- Baker, J. (2011). "The technology-organization-environment framework" In: *Information Systems Theory: Explaining and Predicting Our Digital Society*. Ed. By Y. Dwivedi, M. Wade and S. Schneberger. New York: Springer, 231-246.
- Benlian, A., Hess, T and Buxmann, P. (2009). "Drivers of SaaS-Adoption – An Empirical Study of Different Application Types". *Business & Information Systems Engineering* 1 (5), 357-369.
- Benlian, A. and Hess, T. (2011). "Opportunities and Risks of Software-as-a-Service: Findings from a Survey of IT Executives". *Decision Support Systems* 52 (1), 232-246.
- Berman, S., Kesterson-Townes, L., Marshall, A. and Srivathsa, R. (2012). *The power of cloud - Driving business model innovation*. USA, New York: IBM Institute for Business Value.
- Bhagyashree, A. and Borkar, V. (2012). "Data Mining in Cloud Computing". In: *Proceedings Published by International Journal of Computer Applications*.
- Brynjolfsson, E., Hofmann, P. and Jordan, J. (2010). "Economic and Business Dimensions Cloud Computing and Electricity: Beyond the utility model". *Communications of the ACM* 53 (5), 32-34.
- Etro, F. (2009). "The Economic Impact of Cloud Computing on Business Creation, Employment and Output in Europe". *Review of Business and Economics* 54 (2), 179-208.
- Greene, W. H. (2011). *Econometric Analysis*. New Jersey: Prentice Hall Inc.
- Gujarati, D. N. (2008). *Basic Econometrics*, New York: Mc-Graw Hill Higher Education.
- Gutierrez, A., Boukrami, E. and Lumsden, R. (2015). "Technological, Organisational and Environmental factors influencing managers' decision to adopt cloud computing in the UK". *Journal of Enterprise Information Management* 28 (6).
- Hsu, P., Ray, F. and Li-Hsieh, S. 2014. "Examining Cloud Computing Adoption Intention, Pricing Mechanism and Deployment Model". *International Journal of Information Management* 34 (4), 474-488.
- Landesmann, M.A. (2013). "The new North-South Divide in Europe – can the European convergence model be resuscitated?" *WIIW Monthly Report* 2013 (1), 3-13.
- Low, C., Chen, Y. and Wu, M. (2011). "Understanding the Determinants of Cloud Computing Adoption". *Industrial management & Data Systems* 111 (7), 1006-1023.
- Majava, J., Isoherranen, V. and Kess, P. (2013). "Business Collaboration Concepts and Implications for Companies". *International Journal of Synergy and Research* 2 (1), 23-40.
- Mangula, I.S., Weerd, I. and Brinkkemper, S. (2014). "The Adoption of Software-as-a-Service: An Indonesian Case Study". In: *Proceedings in Pacific Asia Conference on Information Systems (PACIS)*, 2014.
- Marston, S., Li, Z., Brandyopadyay, S., Zhang, J. and Ghalsasi, A. (2011). "Cloud Computing – The Business Perspective". *Decision Support Systems* 51 (1), 176-189.
- NIST. (2009). NIST Definition of cloud computing v15, NIST, Editor. 2009, *National Institute of Standards and Technology*, Gaithersburg, MD.
- Oliveira, T., Thomas, M. and Espadanal, M. (2014). "Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors". *Information & Management* 51 (5), 497-510.
- Rogers, E. (2003). *Diffusion of Innovations*. 5th edition. New York: Free Press.
- Rycroft R. W. (2007). "Does cooperation absorb complexity? Innovation networks and the speed and spread of complex technological innovation". *Technological Forecasting & Social Change*, 74, 565-578.
- Saya, S., Pee, L. and Kankanhalli, A. (2010). "The Impact of Institutional Influences on Perceived Technological Characteristics and Real Options in Cloud Computing Adoption". In: *Proceedings of International Conference on Information Systems (ICIS)*, St. Louis, USA.

- Tornatzky, L. G. and Fleischer, M. (1990). *The Processes of Technological Innovation*. Lexington, MA: Lexington Books.
- Venters, W. and Whitley, E. (2012). “A Critical Review of Cloud Computing: Researching Desires and Reality”. *Journal of Information Technology* 27 (3), 179-197.
- Verma, H. (2013). “Data-warehousing on Cloud Computing”. *International Journal of Advanced Research in Computer Engineering & Technology* 2 (2), 411 – 416.
- Willcocks L., Venters W. and Whitley E. A. (2014). *Moving to the Cloud Corporation*. UK: Palgrave Millam.
- Wu, Y., Cegielski, C.G., Hazen, B.T. and Hall, D.J. (2013). “Cloud Computing in Support of Supply Chain Information System Infrastructure”. *Journal of Supply Chain Management* 49 (3), 25-41.
- Xie X.M., Zeng S.X. and Tam C.M. (2013). “How does cooperative innovation affect innovation performance? Evidence from Chinese firms”. *Technology Analysis & Strategic Management* 25 (8), 939-956.
- Zeng S. X., Xie X. M. and Tam C. M. (2010). “Relationship between cooperation networks and innovation performance of SME”. *Technovation* 30, 181–194.

APPENDIX – Variables definitions – questions

| Variable | Definition |
|-------------------------------------|--|
| Cloud Adoption Propensity | How relevant is cloud computing for your company (very relevant, partly relevant, or not relevant)? |
| IT Investment Reduction Strategy | Has the economic crisis lead you to cancel or significantly downsizing any ICT investments/projects in the last 12 months? (yes/no) |
| Product/Service Innovation | During the past 12 months, has your company launched any new or substantially improved products or services? (yes/no) |
| Process Innovation | During the past 12 months, has your company introduced any new or significantly improved internal processes? (yes/no) |
| Interest in Data Warehousing | Do you consider the topic of data warehouses and data mining to be very relevant, partly relevant, or not relevant for your company? (very relevant/partly relevant/not relevant) |
| Interest in Mobile Services | Do you consider the topic of mobile services such as mobile commerce and remote access technologies to be very relevant, partly relevant, or not relevant for your company? (very relevant/ partly relevant/ not relevant) |
| Innovation Collaboration | Were external experts or business partners involved in developing the new products or services? (yes/no) |
| Electronic Innovation Collaboration | Does your company use online software applications other than e-mail to collaborate with business partners in the development of new products and services? (yes/no) |

A STUDY OF USERS' PERCEPTION OF IT GOVERNANCE DURING INFORMATION TECHNOLOGY ADOPTION IN ORGANISATIONS

Complete Research

Jokonya, Osdan, North-West University, Mafikeng, South Africa, jokonyao@hotmail.com

Abstract

Despite several studies on IT adoption on factors influencing IT adoption in organisations, challenges still persist. This has prompted many researchers to question the suitability of existing models and frameworks for this complex phenomenon in IS research. The existing models and frameworks therefore seem to have limitations in addressing adoption problems faced by many organisations, for example, failing to understand how a framework may facilitate IT adoption in organisations. This paper explores user perceptions on IT governance during IT adoption in organisations. The paper is based on two case studies. The study results indicate most users do not agree with the IT governance process in the two organisations. The results suggest the need to improve the IT governance process in the organisation by involving those to be affected by IT adoption. The study also lays a foundation for further research in this complex phenomenon.

Keywords: IT Governance, case study, IT adoption, technology, framework, model

1. Introduction

Bouwman et al. (2005) defined information technology (IT) adoption as the process of introducing a new technology in the organisation. IT adoption governance is defined as the process that describes how the decision of accepting or rejecting new technology from an individual or organisational context is made. IT adoption governance in an organisation context helps to reconcile different stakeholder demands during the process of adoption regarding the acceptance of information technology. The IT adoption process results in a decision about whether or not to implement the new technology. IT adoption takes places at micro (individual), meso (organisational) and macro (societal) levels (Bouwman et al., 2005).

The challenge of IT adoption in organisations is how to understand its effects at the three different levels. Individual user and organisation effects of technology adoption can never be fully understood because of its complexity and uncertainty. Bouwman et al. (2005) add that technology adoption in an organisation is a two stage process, which involves the organisation and the individual user who decides to use the new technology. The first stage involves the decision by the organisation's executives to embrace a new technology (Rogers; 2003; Bouwman et al., 2005). The second stage involves the individual users in the organisation accepting and using the new technology. The adoption of new technology in an organisation is therefore a two stage process where the organisation makes the decision on a strategic level before the individual users make their own technology adoption decision on an operational level.

Rogers (2003) highlights that individual decision on technology adoption, depends on the organisational decision to a large extent. Bouwman et al. (2005) assert that it is therefore difficult to separate the organisational adoption from the individual decision whether or not to adopt the new technology. The technology adoption in an organisation can be either authoritative (where decisions are made by a few top executives) or collective (where decisions are supported by the members of the social system) (Bouwman et al., 2005). The two types of adoption decision making, authoritative and collective, are suitable for different organisational contexts. Authoritative decision making has been found more suitable for bureaucratic organisations, and collective decision making is more suitable for professional adhocracy organisations (Bouwman et al., 2005).

The size, structure and culture of the organisation are also important in determining the suitable type of IT adoption process (Bouwman et al., 2005). Bouwman et al. (2005) challenge the belief in the existence of a rational decision making process during IT adoption in organisations. The higher level of subjectivity, uncertainty and complexity of the adoption process makes rational decision making difficult in organisations. This study focuses on technology adoption in organisations, which also involves an individual's adoption decision. This paper is structured as follows: section, 2, presents a brief literature on IT Governance. This is followed by section 3 which discusses research methodology. Section 4 then presents the research results. This is followed by section 5 which discusses the results. Finally section 6 of the paper presents the conclusion.

2. Brief Literature Review

The importance of understanding the IT adoption process in organisations has been highlighted by many researchers (Benbasat & Barki, 2007; Lawrence, 2010). Information Technology adoption at individual (for example, a personal mobile device) or organisational level (for example, an ERP system) is a highly researched topic in the IT field. Most literature on IT adoption in organisations is often premised on the assumption that IT adoption benefits are always obvious after selecting the particular technology (Bouwman et al., 2005; Mirza, 2010). However, it has become sufficiently clear that the anticipated benefits from IT adoption in organisations are not self-evident as they may be elusive and difficult to achieve. The pervasive nature of IT and the major dependency of business on IT make it difficult for executives to ignore IT governance in organisations (IT Governance 2011). The risks and benefits associated with IT adoption require it to be governed properly for organisations to reap the IT adoption benefits. Calder (2005) contends that there is no one-size-fits-all solution to IT adoption in organisations.

A study by IT Governance (2011) notes that despite much literature on IT governance frameworks and models, IT problems still persist in many organisations. The study highlights that most of the challenges are to do with issues of people. Calder (2005) argues that adoption of IT in an organisation must critically examine relevance and validity of the organisation specific context. IT Governance (2011) states that stakeholder diversity interests are a major challenge to IT governance in organisations. It suggests the need for a decision framework that may help to alleviate IT adoption problems in organisations. Most existing IT adoption models and frameworks have challenges in meeting stakeholders' needs in organisations (Mirza, 2010).

Goosen et al. (2013) define corporate governance as the relationships and structures that determine the business performance standards and direction. Research on IT governance has shown some inconsistency and lack of consensus on its definition (Goosen et al., 2013). IT governance as part of corporate governance ensures the efficient and effective use of IT resources in organisations (Weill, 2004). IT governance helps to determine which decisions are to be made, and can contribute to the organisation objectives (Turel & Bart, 2014). IT governance is responsible for supporting the effective and ef-

efficient use of IT assets to meet business objectives (Goosen et al., 2013). Some of IT governance definitions include:

- “the organisational structures, processes and mechanisms for decision making and management of IT assets in the organisation” (Goosen et al., 2013).
- “the structure of relationships and processes to control the use of IT in the enterprise in order to achieve the enterprise’s goals by adding value while balancing risk versus return over IT and its processes” (IT Governance, 2011).

According to IT Governance (2011), IT governance focuses on five main areas, which are, value delivery, strategic alignment, performance measure, risk management and resource management. Value delivery ensures that IT investments deliver benefits to the organisation. Strategic alignment ensures that IT investments support business objectives. Performance measure ensures that IT performance in the organisation is reported accurately and timeously based on measurable deliverables and matrices. Risk management ensures that IT related risks are regularly assessed and mitigated in the organisation. Resource management ensures that IT assets are effectively and efficiently deployed in the organisation.

The risk management of IT governance focuses on dealing with issues of risk, compliance and standards using frameworks such as COBIT, ITIL and other ISO standards. Several studies reveal that COBIT benefits are associated with aligning IT with business objectives in organisations (IT Governance 2011). Corporate governance ensures that IT is represented at board level in organisations (IT Governance 2011). IT governance is based on corporate governance principles to manage and use IT to achieve corporate performance goals (Coertze & Von Solms, 2013; Turel & Bart, 2014). Nugroho (2014) says that “IT governance is not concerned with the location and distribution of the IT resources themselves, but rather with the distribution of managerial responsibilities and control that ultimately affect how IT resources are utilized in organisations”.

In the IT decision making area, IT governance focuses on the allocation of decision rights and accountabilities in the use of IT in organisations (Weill, 2004; Musson, 2009; Coertze & Von Solms, 2013; Turel & Bart, 2014). Weill (2004) says “IT governance represents the framework for decision rights and accountabilities to encourage desirable behaviour in the use of IT in organisations. Whereas management is about what specific decisions are to be made, governance is about systematically determining who makes each type of decision, who has input to a decision and how these people are held accountable for their role”.

Amongst other things, IT governance is tasked with deciding on how decision rights and accountability are distributed in organisations to avoid ad hoc decision making (IT Governance 2011). In order to improve IT governance in organisations, Weill (2004) proposes the assignment of decision rights to five IT decision areas (architecture, infrastructure, principle, applications and investment) in organisations. The assignment of responsibilities and roles to decision-making domain areas helps to achieve a balanced governance structure for IT adoption in organisations (Kim et al., 2014).

Kim et al., (2014) note that IT adoption decision rights on organisations are becoming increasingly more important and complex due to diversity of stakeholders needs. Many researchers continue to explore the concept of IT governance in an attempt to find appropriate mechanisms to govern IT adoption in organisations, in order to be accountable to shareholders and stakeholders (IT Governance 2011). Although IT governance as a framework may improve controls with respect to the alignment of IT and business objectives, it pays less attention to how IT adoption decisions are made in organisations (Musson, 2009; Kim & Lee, 2014).

A major challenge for IT governance is the lack of understanding of how decisions are made in order to achieve business objectives (Goosen et al., 2013). Organisations need frameworks to address the IT requirements of different business units' stakeholders in organisations (Weill, 2004); hence IT adoption processes need to involve all stakeholders in organisations (Turel & Bart, 2014). An important factor of an IT governance process is determining where in the organisation decisions are made (Musson, 2009). IT governance needs to promote the participation and engagement of stakeholders in IT adoption issues in organisations (Kim et al., 2014).

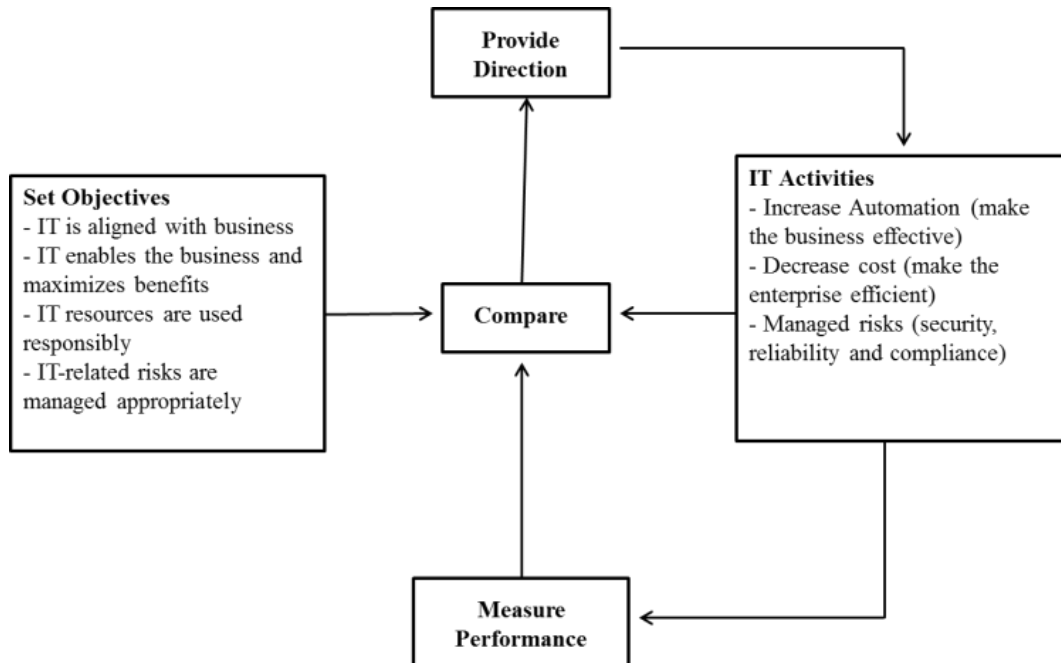


Figure 1. IT Governance Framework (Source: Jokonya 2011:38)

Jokonya (2011) argues that it is important for business leaders (executives) to be fully engaged in IT governance in order for the organisation to achieve business objectives. The biggest challenge in organisations is to reconcile different business units' objectives derived from the nature of their functions. Coertze et al. (2013) note that for an organisation to benefit from IT adoption there is a need for coordination from both IT and the business side of the organisation. The failure to align business and IT strategy will make it difficult to realize value from IT adoption in organisations. IT and business alignment ensures that IT adoption meets the business needs of the organisation (IT Governance, 2011). Turel & Bart (2014) contend that strong IT governance can ensure the proper alignment of business and IT strategy in organisations.

IT Governance Institute (2011) claims that poor IT governance is costing organisations billions of dollars each year after failing to improve business performance as expected. IT governance ensures that IT decisions consider the business objectives and goals as shown in Figure 1. The proper IT governance structure has been credited with better IT related decision in organisations as it helps to get buy-in from business units (Coertze & Von Solms, 2013). The single important predictor of whether an organisation will benefit from IT investment is an effective IT governance structure (IT Governance 2011). Jokonya (2011) suggests that there is a need for new methods to justify IT adoption investment, which identify and quantify the intangible benefits in organisations.

Several researchers found that a balance of IT and business representatives in IT adoption is important to IT governance performance outcomes in organisations (Bowen et al. 2007; IT Governance, 2011).

The same researchers reveal a relationship between project success, business benefits and IT governance. Bowen et al. (2007) point out that the challenge of implementing IT governance in organisations is meeting the needs of multiple stakeholders. Bowen et al. (2007) note that most researchers associate IT governance with the structure for making IT decisions in organisations. Bowen et al. (2007) contend that focusing on the structure aspect only of IT governance ignores other important activities. Jokonya (2011) contends while organisations appreciate the importance of IT governance they realize that it is a complex undertaking as it is difficult to implement. De Haes and Van Grembergen (2008) propose an IT governance framework based on structure (defined roles and responsibilities), processes (strategic decision making, standards and frameworks) and relational mechanisms (dialogue between IT, Stakeholders and business).

Several researchers concede that the challenge of IT adoption is getting agreement from different stakeholders' constituencies who disagree on goals and how to achieve them (Cordoba, 2009, Coertze & Von Solms, 2013). The reductionist approach to IT governance has been seen to pose challenges to complex modern organisations (Jackson, 2010, Cordoba, 2009). The complexities of IT governance mean organisations need to work with different paradigms to offer multiple insights to the complex phenomenon (Turel & Bart, 2014). The systems approach has been credited with acknowledging the complexity and dynamic nature of IT governance in organisations (Cordoba, 2009).

In summary, whilst the literature acknowledges the importance of effective IT governance in organisations, it is still a challenge to many organisations. The major challenge is satisfying different stakeholder constituencies with different worldviews in organisations, to the benefits of IT adoption. The challenges of IT adoption in organisations may require different paradigms to understand the problem context before making decisions. The next section discusses the research methodology.

3. Research Methodology

The case study research strategy in information systems research is accepted because of its suitability to provide understanding of the relationship between organisations and technology (Oates, 2009). Case study research has been found to be a viable research strategy in information systems because of its in-depth approach. The case study research provides an opportunity for the researcher to understand IT governance in a normally inaccessible phenomenon (Oates, 2009). The case study also means that the participants are exposed to the same IT adoption environment, which is useful to understand their perceptions from a single setting.

Most of the contemporary issues dealt with in case study research are common to many organisations. In addition case study research provides an opportunity of first hand evidence on a particular phenomenon. Case study research may adopt a single case or multiple case designs depending on the research objectives (Eisenhardt et al., 2007). One of the contentious questions has been how many cases are sufficient for multiple case studies and there is no simple answer as it depends on the research purpose and question (Rowley, 2002). This research used two case studies, which were sufficient for the research objectives.

3.1 Case study selection

The research site was purposefully selected to help answer the research questions. The researcher used purposive sampling in order to select data collection units that yielded the most relevant and broad range of perspective and information of the research area (Yin, 2009). Two companies agreed to participate in the study from six that were approached by the researcher. The basis for selecting the two

organisations for the research was the diversity and appropriateness of cases. Data was gathered about participants' opinions on the importance of IT Governance during on IT adoption in organisations.

3.2 Data collection

The data collection, using questionnaires, took three months. About two hundred questionnaires were distributed to employees of the two companies. The questionnaires had pre-defined questions seeking participants' perceptions. A five point Likert scale was used to develop the questionnaire. The questionnaire was pre-tested with a few participants to refine the questions.

The questionnaire variables were tested using Cronbach's alpha values to see if they were reliable and acceptable. The questionnaire variables were above 0.80, which is an indication of good reliability at 0.88 (Van Voorhus et al., 2007). About ninety valid questionnaires were returned from the two companies. The returned questionnaires represented a 45 percent response rate. Quantitative data from the questionnaires was captured and analysed using SPSS (Statistical Package for the Social Science) version 21.

3.3 Sample size

Quantitative research offers guidelines on sample sizes needed for different statistical procedures, unlike qualitative research which does not have an agreed sample size (Nunnally, 1978). The sample size of 90 cases was therefore found adequate for the required statistical procedures. The descriptive statistics used include frequency tables, means, T-test and analysis of variance (ANOVA) to provide summarized data for discovering trends, patterns and ease of communication and understanding.

4. Results

This section presents the results from the data collected using questionnaires. The organisations that participated in the study were all using enterprise resource planning (ERP) systems. The selected participants were previously involved in IT adoption in their organisations. The study was, therefore, interested in their perceptions based on previous experience in IT adoption in their organisations. The IT governance construct measures the respondents' perception on the governance process during IT adoption in their organisations. The IT governance variables were adapted from the literature. The questionnaire was pretested as part of refining some of the unclear questions. The variables of IT governance were tested to see their association. This section is organized as follows: section 4.1 presents the demographic data, section 4.2 presents the frequencies of the variables, section 4.3 presents t-test results, and finally section 4.4 presents the analysis of variance results.

4.1 Demographic results

This section presents the respondents' demographic characteristics which were: company type, age, gender, departments, position, education, involvement in IT adoption, number of years in the organisation, and member status. There were equal numbers of respondents from retail and manufacturing companies. In terms of age, 30 percent of the respondents were 30 years and below, 37 percent were between 31 and 40 years, 26 percent were between 41 and 50 years and finally 8 percent were 51 years and older. The majority of the respondents were males - 54 percent - compared to 46 percent for female respondents.

About 24 percent of the respondents were from the finance department, 20 percent from information technology and 66 percent from other departments. In terms of positions of respondents, 28 percent

were clerical, 27 percent were managers, 9 percent were supervisors and 37 percent were from other positions in their organisations. About 42 percent of the respondents had matriculation as their level of education compared to 36 percent of the respondents who had a first degree. The remaining 22 percent of the respondents had a second degree as their level of education.

About 42 percent of the respondents were involved in IT adoption in their organisation compared to 58 percent of respondents who were not involved. In terms of number of years in the organisation, 34 percent of the respondents had two years or less, 26 percent had between three and five years, 20 percent had between six and ten years and 17 percent had more than ten years in the organisation. About half of the respondents were members of a committee in their organisations whilst half were not members of any committee in their organisations.

4.2 IT governance frequencies

Figure 2 below shows frequencies of respondents on the IT governance construct variables. Only a third of the respondents agreed that IT adoption was always transparent compared to those who disagreed or were neutral. Also about a third of the respondents agreed that IT and business objectives always aligned compared to those who disagreed or were neutral. About less than a quarter of the respondents agreed that objectives were always clear to stakeholders compared to those who disagreed or were neutral. Slightly less than a third of the respondents agreed that there was shared responsibility in IT adoption compared to those who disagreed or were neutral. Slightly above a quarter of the respondents agreed that the desired outcomes were always clear compared to those who disagreed or were neutral.

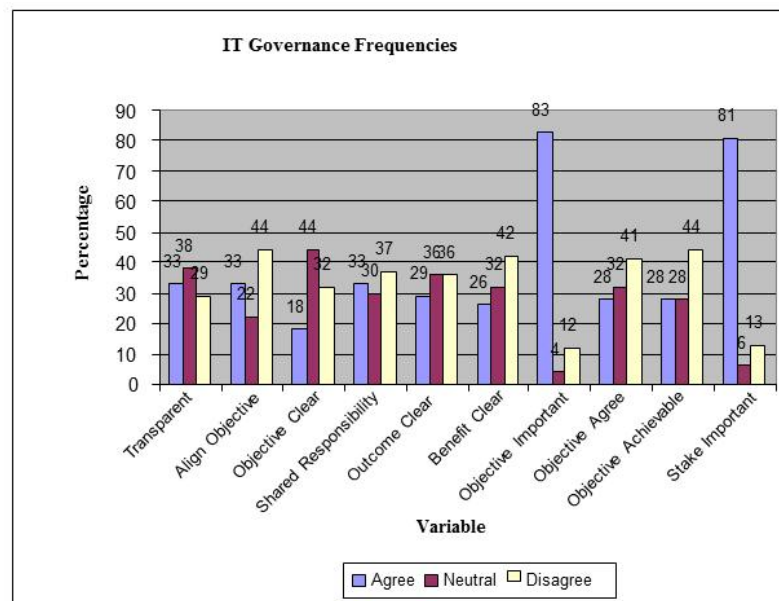


Figure 2. IT Governance Frequencies

Slightly above a quarter of the respondents agreed that benefits were always clear compared to those who disagreed or were neutral. A majority of the respondents, 44 percent, agreed that clear IT objectives were important compared with a few who disagreed or were neutral. Slightly above a quarter of

the respondents agreed that there was always agreement on IT adoption objectives compared to those who disagreed or were neutral. Slightly above a quarter of the respondents agreed that IT adoption objectives were always achieved compared to those who disagreed or were neutral. A majority of the respondents, 81 percent, agreed that stakeholder participation was important compared to those who disagreed or were neutral. The results show only two of IT governance constructs variables that had more than fifty percent in agreement. The respondents agreed that IT objectives and stakeholder involvement were important during IT adoption in organisations. The few who agreed suggested that respondents were not pleased with IT governance in their organisations. The results suggest the need for improvement in IT governance in the organisations. In summary, the results suggest that the organisations have challenges with IT governance issues. The next section presents the t-test results between demographic variables with two categories and IT governance variables.

4.3 T-Test results of demographic variables

The t-test was useful to assess significant differences between the mean constructs of demographic variables with two categories such as company, sex and so forth (Table 1). The T-test conducted showed significant differences between the two companies on users' perception on shared responsibility during IT adoption (.015) and benefits of IT adoption are always clear to stakeholders (.008). The gender demographic variable also showed a significant difference on respondents' perception on shared responsibility during IT adoption in organisations (.020) and the importance of clear IT adoption objectives in organisations (.002). The involved user demographic variable showed significant differences on the importance of clear IT adoption objectives in organisations (.000), the existence of agreements on IT adoption objectives (0.011) and the importance of stakeholder participation during IT adoption in organisations (.007). The committee member demographic variable showed significant differences on the existence of clear IT adoption objectives to stakeholders (.002) and the importance of clear IT adoption objectives in organisations (.000).

| Independent Variable | Dependent Variable | F Value | Sig. | Categories |
|----------------------|--------------------------------|---------|----------|--------------------------|
| Company | Shared responsibility | 6.16 | 0.015* | Retail and manufacturing |
| Company | Clear benefits | 7.35 | 0.008** | Retail and manufacturing |
| Gender | Shared responsibility | 5.64 | 0.020* | Male or Female |
| Gender | Clear objectives | 9.82 | 0.002** | Male or Female |
| Involved | Importance of clear objectives | 19.18 | 0.000*** | Yes or No |
| Involved | Agreements on objectives | 6.81 | 0.011* | Yes or No |
| Involved | Stakeholder participation | 7.74 | 0.007** | Yes or No |
| Member | Existence of clear objectives | 19.16 | 0.002** | Yes or No |
| Member | Importance of clear objectives | 7.38 | 0.000*** | Yes or No |
| Company | IT Governance score | 2.11 | 0.038* | Retail and manufacturing |

Table 1. T-test for demographic variables, Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, (n=90)

The t-test conducted showed significant differences on IT governance ($p=0.038$) between the two companies' retail and manufacturing respondents in terms of perceptions. The results suggest that there are

significant differences between the respondents from the two companies with regards to their perception on IT governance in their organisations. The results show that manufacturing respondents agreed more on the IT governance construct than did retail respondents (Table 1). The next section presents the analysis of variance between demographic variables with more than two categories and IT governance variables.

4.4 Analysis of variance of demographic variables

The analysis of variance (ANOVA) was used to assess significant differences between demographic variables with more than two categories. In addition multiple post hoc was used to reveal which among the more than two categories differed significantly from one another (Table 2).

| Independent Variable | Dependent Variable | F-Value | Sig. |
|----------------------|--------------------------------|---------|---------|
| Department | Objectives always clear | 2.38 | 0.046* |
| Department | Importance of Clear Objectives | 4.46 | 0.001** |
| Education | Objectives always Achieved | 4.50 | 0.032* |
| Education | IT Governance Score | 3.78 | 0.028* |

Table 2. ANOVA for demographic variables Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ($n=90$)

The analysis of variance (ANOVA) conducted on the department demographic variable showed a significant difference on the existence of clear IT adoption objectives to stakeholders (.046) and the importance of clear IT adoption objectives in organisations (.001). The education demographic variable showed differences on whether IT adoption objectives are always achieved in organisations (.032). The analysis of variance was conducted between people’s education levels and perceptions of IT governance (Table 2). The results suggest that education level had an influence on perceptions about IT governance ($p < .028$). Less educated matriculated respondents were more in agreement with the IT governance than first and second degree respondents. The more educated they were the less likely were respondents to agree with the IT governance. The correlation analysis conduct between demographic variables and IT Governance variables did not show any significant correlation. Having discussed analysis of variance and correlation analysis, the next section presents the discussion of the study.

5. Discussion

The results of the study show that in general most respondents do not agree with the IT governance process in the two organisations as show by statistics (Figure 2). Most of the respondents disagreed with most the variables of IT governance except the two: objectives are important during IT adoption and stakeholders are important during IT adoption. The issues are very important because they help to get buy-in from those to be affected by the IT adoption process. The buy-in is important to have positive outcomes during IT adoption in organisations.

The t-test conducted showed significant differences between demographic variables and four IT governance construct variables with two categories which support the frequency results that there are disagreements on IT governance issues in the two organisations. The demographic variables, which

showed significant differences, include company, gender, involved and committee members. Most of the IT Governance variables showed significant differences in the two organisations which is of major concern to the success of IT adoption in the organisations. This shows that the organisations have challenges with IT governance as most participants did not agree on most of the IT governance variables.

The analysis of variance also supported the frequency and t-test results with regard to significant differences between respondents' perception on IT governance process in the two organisations. Most of the significant differences were based on department and education demographic variables. The analysis of variance results suggest that education level had an influence on perceptions about IT governance process in the organisations. Less educated matriculated respondents had a more positive perception than the more educated respondents with regard to IT governance process in the organisations. This is also the same with department, with some more positive than others with regard to IT governance processes in the two organisations. There was no correlation between nominal demographic variables and IT governance construct variables.

The results suggest that users' perceptions during IT adoption in organisations may be influenced by their involvement and having clear objection during the process as shown by two variables with highest percentage in Figure 2. The results suggest the need for improvement in IT governance processes in the organisations. In summary the results suggest that the organisations have challenges with IT governance processes which call for improvement. Having discussed the results the next section presents the conclusions of the study.

6. Conclusion

The study suggests that it may be worthwhile to consider engaging stakeholders during IT adoption in organisations to improve the governance process and influence positive outcomes. The stakeholders engagement will help to reach consensus, secure buy-in and accommodate different worldviews during IT adoption in organisations. In addition, the study suggests that stakeholder engagement may be important in changing the negative attitudes of users during IT adoption in organisations. The study contributes to our understanding of users' perceptions of IT governance during IT adoption in organisations. In addition, the study adds to our understanding of IT adoption challenges in organisations.

Although the study contributed to the understanding of IT adoption in organisations it has its limitations which need to be acknowledged. One of the major limitations of the study is that it is based on case study research which makes it difficult to generalize the results of the study. However, the limitation provides an opportunity for further research using a survey which can allow the results to be generalized to a large population. In addition, there is a possibility to employ other advanced statistical methods such as factor analysis and structural equation modelling. This study therefore acts as a stimulus and provides avenues for several areas of further research on this topic of IT Governance in organisations.

References

- Benbasat, I. and H. Barki (2007). "Quo vadis TAM?" *Journal of the Association for Information Systems*, 8(4), 211-218.
- Bowen, P. L.; M. Y. D. Cheung and F. H. Rohde (2007). "Enhancing IT governance Practices: A Model and Case Study of an Organisation's Effort". *International Journal of Accounting Information Systems*, July 2007, 8(2007), 191-221.

- Bouwman, H.; J. Van Dijk; B. Van den Hoof and L. Van de Wijngaert (2005). *Information and Communication Technology in Organisations: adoption, implementation, use and effects*. London: Sage.
- Calder, A. (2005). *IT Governance: Guidelines for Directors*, IT Governance Ltd.
- Coertze, J. and Von Solms, R. (2013) "The Board and IT Governance: A Replicative Study". *African Journal of Business Management*, 7(35), 3358-3373.
- Cordoba, J. R., (2009). "Critical reflection in planning information systems: A contribution from critical systems thinking", *Information Systems Journal*, (19.2) 123–147.doi:10.1111/j:1365-2575.2007.00284.x
- De Haes, S. and W. Van Grembergen (2008). "Analysing the Relationship between IT Governance and Business/IT Alignment Maturity". *Proceedings of the 41st Annual Hawaii International Conference System Science*. Waikoloa, HI, 428-428.
- Eisenhardt, K. M. and M. E. Graebner (2007). "Theory Building from Cases: Opportunities and challenges", *Academic of Management Journal*, Vol 50 No. 1 pp. 25-32,
- Goosen, R. and R. Rudman (2013). "An Integrated Framework to Implement IT Governance Principles at a Strategic and Operational Level for Medium to Large-Sized South African Businesses". *International Business and Economics Research Journal*, July 2013, 12(7), 835-854.
- IT Governance Institute (2011). Board Briefing on IT Governance: Rolling Meadows, IL. *IT Governance Institute Ltd*.
- Jokonya, O. (2011). *Information Technology Governance to Achieve Business Objectives: Information Technology Governance as a Tool for Achieving Business Objectives*. Saarbrucken: LAP LAMBERT Academic Publishing GmbH & Co. KG.
- Jokonya, P. (2014). "A framework to assist organisations with information technology adoption governance" PhD thesis. University of South Africa.
- Jackson, M. C. (2010). "Reflections on the Development and contributions of Critical Systems Thinking and Practice". *System Research and Behavioural Science* 27, 133-139
- Kim, K. and H. Lee (2014). "ITA Functions and IT Governance from Towards Public and Private Enterprise in Korea: A Study for Influence Factors". *Journal of Scientific and Industrial Research*, 73(2014), 16-20.
- Martins, L. M.; A. Moura; P. R. D. Cunha and A. D. D. Figueiredo (2010). "Selecting and Ranking IT Governance Practices for Electric Utilities". *Proceedings of the Americas Conference on Information Systems*, Lima, Peru.
- Mirza, A. A. (2010). "Critical Success and Failure Factors in a Charitable Organisation's Systems Development and Adoption". *Proceedings of the Americas Conference on Information Systems*, Lima, Peru, Paper 381.
- Musson, D. (2009). "IT Governance: A Critical Review of the Literature. In: CATER-STEEL, A. (ED). *Information Technology Governance and Service Management: frameworks and adaptations*". (63-80), Hershey, PA: IGI Global.
- Nugroho, H. (2014). "Conceptual Model of IT Governance for Higher Education based on COBIT 5 Framework". *Journal of Theoretical and Applied Information Technology*, February 2014, 60(2), 216-221.
- Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill
- Oates, B. J. (2009). *Researching Information Systems and Computing*, London: Sage
- Rogers, E. M. (2003). *Diffusion of innovations*. 5th ed. New York: The Free Press.
- Rowley, J. (2002). "Using Case Studies in Research", *Management Research News*. Vol. 25 No. 1,
- Turel, O. and C. Bart (2013) "Board-level IT Governance and Organisational Performance". *European Journal of Information Systems*, 23(2013), 223–239.
- Van Voorhis, R. W. C. and L. B. Morgan (2007). "Understanding Power and Rules of Thumb for determining Sample sizes", *Tutorials in Quantitative Methods for Psychology*, Vol. 3(2), pp. 43-50,
- Weill, P. (2004). "Don't Just Lead, Govern: How Top- Performing Firms Govern IT". *MIS Quarterly Executive*, ta, 3(1), 1-17.

Yin, R. K. (2009). *Case study research, Design and methods*. 4th. ed. London: Sage

RATIONALITY VS FASHION IN CLOUD ADOPTION DECISIONS: THE CASE OF CLOUD-ENABLED PAYROLL SYSTEMS

Completed Research

Polyviou, Ariana, Athens University of Economics and Business, SingularLogic S.A., Athens, Greece, apolyviou@aueb.gr

Pramatari, Katerina, Athens University of Economics and Business, Athens, Greece, k.pramatari@aueb.gr

Pouloudi, Nancy, Athens University of Economics and Business, Athens, Greece, pouloudi@aueb.gr

Abstract

Cloud computing is an innovation said to be bringing radical changes in the way organisations interact with information systems. Decision makers seek for reasons that make the adoption of cloud computing meaningful (or not) for their organisation. At the same time, several stakeholders strive to promote cloud as a ‘must-do’ innovation. Given the hype that currently surrounds cloud computing, it is worth investigating whether the rationale behind the decision to adopt cloud is partly muddled by an apparent simplicity of choice for decision makers: ‘to adopt or fail’. Based on indications in the existing literature and recent empirical findings, we argue that in addition to the factors grouped under cloud’s relative advantage, IT-Fashion also influences the intention to adopt cloud which is currently overlooked. By using the example of cloud-enabled payroll systems, we demonstrate that IT-Fashion has also a significant impact on the intention to adopt cloud services.

Keywords: Cloud adoption, IT-Fashion, Rational factors, payroll systems.

1 Introduction

Managers often come across innovation waves, each of them promising to be providing remarkable advantages to the organisation if adopted (Swanson 2012). To decide whether a new technological innovation is worth adopting, managers need to make sense of the innovation in the context of their organisation. This decision making process is rational to the extent that the consequences of an innovation's adoption are understood (Rogers 1995). However, during the first half of an innovation's diffusion curve, consequences of an innovation cannot be fully inferred since the innovation has not been widely adopted yet. Seeing this as an opportunity, fashion setters (e.g. vendors, consultants, IT and management 'gurus'), selectively promote certain innovations creating buzz around them and using 'must-have' rhetoric (Abrahamson 1991). In this way, selected innovations generate a bandwagon phenomenon and colonize practice (Wang 2010). As a result an organization's decision for innovation adoption at this stage, may not be based only on purely rational criteria, but is also likely to be influenced by fashion (Swanson et al. 1997). Thus, the reasoning associated with the decision to adopt the technological innovation is muddled by a dilemma: 'to adopt or fail' (Kieser 1997).

Cloud computing is such a technological innovation, said to be radically changing the way in which organisations interact with IT. It is deemed to be transferring the world from the PC era to the utility computing era (Carr 2009). Compared to previous technologies, the use of cloud computing minimizes IT investment risks, offers computing power which is tailored to the demand that each organization has at each moment, provides the ability to access technological resources from everywhere, increases the potential of the organization to innovate and alters the organization's boundaries with its employees, customers and other organizations (Köhler et al. 2010; Willcocks et al. 2014). In this paper, cloud computing is defined as a form of shared-resource and elastic computing services which are offered on-demand and are pooled via external data-centres and are made available via the Internet (Bohn et al. 2011; Durkee 2010). Cloud comprises three different layers, Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS).

IS scholars have already taken steps to understand cloud adoption decisions. While early work focused on enlisting the advantages and disadvantages of cloud computing (Janssen et al. 2011; Köhler et al. 2010), recent studies have employed more sophisticated approaches by employing widely-cited IS adoption literature to frame their studies (Low et al. 2011; Oliveira et al. 2014). Their findings highlight, inter alia, that an important factor influencing the adoption of cloud services is its relative advantage; the extent to which cloud computing brings advantages to the organisation compared to previous technologies used (Rogers 1995). At the same time, current studies indicate that, beyond relative advantage, factors of psychological nature (i.e., attitude towards the term cloud) seem also to have a remarkable influence on cloud adoption decisions and merit further research (Morgan et al. 2013). Despite these indications, cloud adoption literature has not addressed the impact of IT-Fashion so far.

In this paper we investigate whether, in addition to the 'rational' factors that influence cloud computing adoption (i.e. cloud's relative advantage), IT-Fashion also influences the intention to adopt. Drawing on diffusion of innovation and fashion management theories, we formulate a theoretical framework that takes into account both rational and IT-Fashion factors. The objective of our study is to juxtapose their influence. To achieve this, the study draws on the results of a previous preliminary research, specifying the factors that express the relative advantage of cloud computing (Polyviou et al. 2014a). We empirically test the validity of our hypotheses through a survey investigating the adoption of cloud-enabled payroll systems (i.e., payroll SaaS systems). The findings suggest that IT-Fashion has an influence on the intention to adopt cloud. Thus, this paper contributes to cloud computing research, offering a complementary lens for understanding cloud adoption decisions.

The paper is structured as follows. The following section reviews the relevant literature and presents our theoretical model. Section 3 defines the methodology employed to test our hypotheses. Section 4

describes the results of the empirical study, while Section 5 discusses their implications for IS research. Section 6 presents the study's conclusions and proposes further research directions.

2 Related Literature and Hypotheses Development

2.1 Theoretical Background

Rogers's (1995) Diffusion of Innovation (DOI) theory defines innovation as an idea, practice or technology that is new to an organisation. According to DOI, once an innovation is released, it is spread to the targeted audience through several channels (e.g. mass-media) across time. In this way, managers can be informed about the innovation and be motivated to figure out further information about it so that the organisation can create an attitude towards the innovation (Rogers, 1995). In order to formulate and finalise their attitude towards the innovation (favourable or unfavourable), it is necessary for managers responsible for the adoption decision to find reasons for (or against) the innovation's adoption. In other words, in this stage, decision makers construct the rationale leading them to decide whether to adopt or reject the innovation. This is a difficult process, especially during the first half of the diffusion innovation's curve, as the innovation has not been widely adopted yet, so that its consequences become clear. Hence, at this stage the decision involves a large amount of uncertainty. One of the characteristics of the innovation influencing the decision for or against its adoption according to Rogers (1995) is its relative advantage. Rogers (1995: 212) defines relative advantage as the degree to which the innovation is perceived as a better idea compared to the one it supersedes.

To complement our understanding of how decision makers formulate their attitudes towards an innovation, we draw on fashion management theory. Fashion management theory complements diffusion of innovation theory as it emphasises the social setting of the innovation. Fashion is defined as a relatively transitory belief that a technique is fresh, efficient and lays a forefront of management practice (Abrahamson 1991). According to Abrahamson (1996), attitude towards innovation is influenced by the adopter's social norms and pressures. Hence, fashion management encourages researchers to take into account fashion setters when examining the diffusion of an innovation (Abrahamson 1996). Fashion setters, as for example consultants, academics, journalists and gurus, select and promote certain innovations as being at the forefront of management practice, using arguments arising by the anecdotal success stories of early adopters. Fashion setters are said to deploy media as means to disseminate certain views and drive the formulation of perceptions surrounding an innovation. In this way, the rationale behind the adoption of the innovation decision is transformed into an oversimplified dilemma for the decision maker: to 'adopt or fail' (Kieser 1997). As the adoption rate rises, the beliefs about the innovation generated by the fashion setters become stronger. This generates a collective belief towards the innovation. Meanwhile, it formulates a self-reinforcing cycle through which innovation diffusion and fashion management levels build on each other and thus they develop a bandwagon phenomenon for the particular innovation (Wang 2010).

2.2 Cloud Adoption Literature

Recent research work steps on understanding cloud adoption decisions, focuses on identifying the advantages and disadvantages of cloud computing or exploring the consumer preferences on certain cloud layers (Anandasivam et al. 2010; Giessmann et al. 2012; Janssen et al. 2011; Polyviou et al. 2014a). As the focus of this work is on the technology per se, it does not account for non-technological factors influencing cloud adoption (e.g. external factors). Responding to this challenge, other cloud adoption scholars employ widely-cited IS adoption theories such as the Technology-Organisation-Environment framework (TOE framework) (Tornatzky et al. 1990), the Human-Organisation-Technology model (HOT-fit model) (Yusof et al. 2007) and Diffusion of Innovation theory (DOI) (Rogers 1995) to guide their study using respondents from selected industries or sectors (Alshamaila et al. 2013; Low et al. 2011; Morgan et al. 2013; Oliveira et al. 2014). Such studies identify several fac-

tors that are influencing, at least to some extent, cloud adoption e.g. relative advantage, technological readiness, top management support, legal issues, complexity, competitive pressure.

Closer to our work is research that proposes a theoretical background that combines diffusion of innovation with other theories. For example, Oliveira et al. (2014) employ some of the characteristics of innovation defined by Rogers (1995) in combination with the TOE framework to test the impact of a set of factors on the decision to adopt cloud services (Oliveira et al. 2014). Similarly, Morgan & Conboy, (2013), use DOI theory and TOE framework to conduct a qualitative exploratory study on cloud adoption factors (Morgan et al. 2013). Both studies, evaluate factors grouped under relative advantage, as highly influential for cloud computing adoption. This conclusion is also confirmed by other studies examining cloud adoption factors (Alshamaila et al. 2013; Low et al. 2011).

The use of traditional IS adoption theories has been adequate for understanding the adoption of IS innovations in the past (Chau et al. 1997; Kuan et al. 2001). Undoubtedly, it has help us to reveal that relative advantage is one of the dominating factors influencing cloud computing. However, recent research efforts on cloud computing adoption reveal there are insights that cannot be sufficiently explained by the theoretical frameworks employed so far. More specifically, findings of Morgan and Conboy (2013) underline that factors influencing cloud computing adoption are not only technological, but also psychological. By psychological, authors refer to the perception of the decision makers towards the term ‘cloud’ itself, which seems to be also a factor influencing the decision to adopt. To this end, the authors call for further research in this direction, as the theoretical lenses employed by their study do not provide sufficient explanation (Morgan et al. 2013). At the same time, the pertinent literature also suggests that cloud may be currently passing through a ‘fashion hyperbole’ (Whitley et al. 2013). Building on these suggestions, we propose the use of Fashion Management theory as a complementary lens for addressing the factors that affect cloud computing adoption.

Thus, the research question tackled by this paper is formulated as follows:

To what extent is IT-Fashion influencing cloud computing adoption compared to the factors related to cloud’s relative advantage?

The following subsection presents a set of hypotheses formulated to address this research question.

2.3 Hypotheses Development

As underlined by recent literature findings, the relative advantage of cloud computing compared to previous technologies is what drives the adoption intention the most. Some adoption researchers attempt to interpret relative advantage in the context of cloud computing. For example, Morgan and Conboy (2013) identify cost reduction as one of the factors leading to cloud’s relative advantage. Similarly, in their study Oliveira et al. (2014) break down cloud’s relative advantage into cost reduction. Using related studies in combination with the findings of our preliminary study results, we attempt to analyse the relative advantage characteristic as defined by Rogers (1995) in terms of factors that define cloud’s advantage compared to previous technologies. To enhance our knowledge on cloud computing adoption decision, a preliminary qualitative study conducted through semi-structured interviews with key stakeholders was carried out as described in (Polyviou et al. 2014a). The interview agenda aimed at capturing the factors that affect cloud computing adoption. Open-ended questions were included allowing the participants to spontaneously present their views. Among the factors identified, outstanding were the factors relevant cloud’s technology advantage compared to what it supersedes. In addition, indications about IT-Fashion could also be identified.

Firstly, most of the interviewees underlined the cost reduction as one of the main factors leading to cloud adoption. This is relevant to the reduction of costs concerning the investment in IT, IT maintenance, upfront investment and capital expenditure for IT. Cost reduction, has also been identified by existing literature as an important factor for cloud adoption (Janssen et al. 2011; Polyviou et al. 2014b). This finding was also supported by the interviews conducted in the frame of our preliminary

study. Hence, we consider cost reduction as one of the factors expressing cloud's relative advantage and based on this we define the first hypothesis guiding our research:

H1: Perceived cost reduction positively influences intention to adopt cloud services

Furthermore, access from everywhere was identified as one of the key factors driving cloud adoption. As denoted by one of our interviewees, "it is accessible from anywhere. Hence employees could even work from home without being obliged to stay in their offices"[quote of chief information officer in Poland]. Similarly, the cloud adoption literature recognises 'access from everywhere', as a factor influencing the decision to deploy cloud services for an organisation (Chau et al. 1997; Janssen et al. 2011; Köhler et al. 2010). Based on the above, we consider that remote access a functionality that cloud technology offers which was not provided by any previous technology used and thus it can be expressed as cloud's relative advantage. This defines the second hypothesis of our study:

H2: Perceived remote access positively influences intention to adopt cloud services

In addition, previous technologies tended to incorporate the storage or processing of the organisation's data within the premises of the organisation. In contrast, the deployment of cloud services implies that the data of the organisation is stored, processed and handled by an external provider in a multi-user environment (Joint et al. 2009). This may generate additional concerns to the organisation concerning the loss or exposure of sensitive data and thus it may act as a barrier towards cloud adoption (Janssen et al. 2011; Köhler et al. 2010; Oliveira et al. 2014; Polyviou et al. 2014b). Similarly, the results of our preliminary study indicate that security concerns is an important factor negatively influencing cloud adoption. As noted by one of the interviewees "possibility of leaking of sensitive data [quote of chief information officer in Greece]" is one of the main factors negatively influencing cloud adoption. Based on the above we extract the third hypothesis of our research:

H3: Security concerns negatively influence intention to adopt cloud services

Moreover, Morgan and Conboy (2013) underline that, beyond technical factors, psychological factors also have an important influence on the adoption decision. According to the vendors included in our preliminary study, the decision to move to the cloud is believed to be difficult for the decision makers, as they struggle to minimise the degree of uncertainty involved in their decision. At the same time, vendors note that they use the term 'cloud' in their campaigns as this creates positive associations in most of the cases, although they acknowledge that decision makers may not have a full understanding of the term: "We used the term cloud [...] because that's really positive [...] We believe that many people come from internet search [...] -Do you think that people who search for 'cloud' understand what they are searching for? No! [quote of cloud vendor in Germany]". Also, vendors of our preliminary study also denoted that they often exhibit their products through various events, fares and conferences promoting best practices in certain sectors. To this end, the fourth hypothesis of our study was derived:

H4: Tendency to follow IT-Fashion positively influences the intention to adopt cloud services

Based on the hypotheses defined in this section, we design the model of our study as illustrated in Figure 1. As we consider the factors grouped under the relative advantage characteristic to be defining the 'cloud-enableness' of the technology under consideration, our dependent variable is defined as 'intention to adopt cloud-enabled systems'. In the section that follows, the use case employed to test our model as well methodology followed for our study are being discussed.

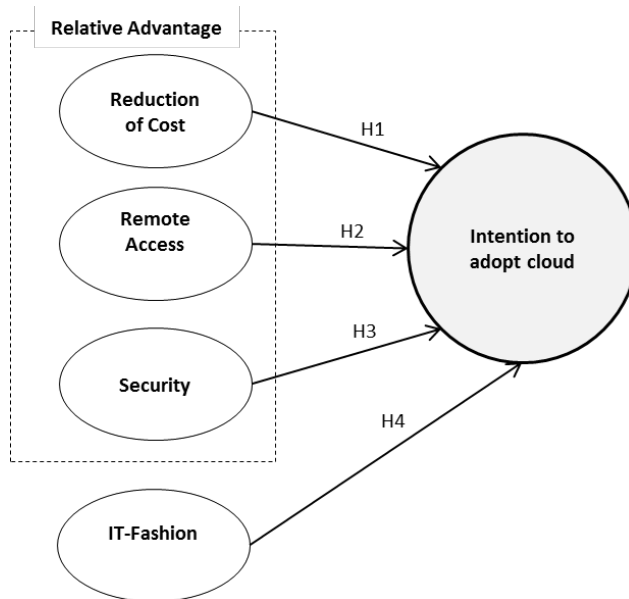


Figure 1. Hypotheses Model

3 Research Methodology

To investigate the factors that influence a firm’s decision to adopt cloud services and test the hypotheses discussed in the previous section of this paper, a quantitative methodology was followed. A survey was deployed using the firm as the unit of analysis. The questions targeted the case of payroll systems and, in particular, the adoption of cloud-enabled payroll systems. We chose to focus on payroll systems, as there is a current trend among software providers in this market to upgrade their solution and move from on-premise software installations to cloud-enabled payroll systems, following the software-as-a-service model. This trend is also met in other functionality domains often related to back-office operations that are supported over the years by mainstream software solutions. Moreover, the fact that payroll systems belong to the group of non-strategic and infrequently-changed systems that a firm owns would suggest that decisions regarding the use of these systems is even less influenced by IT-Fashion.

The respondents targeted were mainly the managers of the firms’ HR department or other senior members of the HR department. In few cases, e.g. in cases that there was no HR executive directly related to the payroll management of the firm, the targeted audience was the CEO/Director or the CIO of the firm. Our sample included Greek firms of several sizes and business sectors. The survey was made available through the use of an online survey tool (Qualtrics). To collect the dataset, an initial email inviting participants to respond to the survey was distributed and a reminding email, as well as follow-up phone calls, were used as means to remind participants to respond. To encourage participation respondents were informed that, if interested, they could receive the results after the end of the study. Each company was asked to respond to the questionnaire only once. Responses were collected within the interval of 16 days in the autumn of 2014. Data provided by firms that have their payroll procedure outsourced, were excluded from the dataset considered for this paper. Out of the 682 emails, a total of 69 valid questionnaires were collected (the 29 additional responses by companies outsourc-

ing their payroll were excluded). The response rate of the respondents invited to participate in the study was 10% which is typical for studies targeting senior managers via emailed survey.

3.1 Measurement & Scale development

To address the impact of certain factors on the intention to adopt cloud-enabled payroll systems, our model includes only first-order constructs capturing both the rationality and the IT-Fashion concepts introduced in the previous section. The cost reduction, remote access desire and security concerns constructs are employed for instantiating the impact of rationality on the intention to adopt cloud-enabled payroll systems. The IT-Fashion construct is used for measuring the impact of IT-Fashion. Finally, the model captured three control variables; Industry sector, HR department existence and Number of employees. The constructs and items included in the survey are enlisted in Appendix A.

In most of the cases, items were adapted from existing literature, whereas in cases where the desired construct measurements were not available in existing literature, items were self-developed. The intention to adopt a cloud-enabled payroll system might be influenced by the desire to reduce costs. Thus, cost reduction in the proposed model was measured using three items and the responses were provided in a seven-point Likert-type scale with endpoints Strongly Agree (1), Strongly Disagree (7) (Premkumar et al. 1999). Another advantage offered by cloud-enabled payroll systems is the ability to access the system remotely from anyplace, anytime as this might improve the quality of work regarding the payroll task. To this end, the desire for remote access functionality for the payroll system was measured through the 'remote access' construct. The construct measured by three items with answers provided in seven-point Likert-type scale with endpoints Strongly Disagree (1), Strongly Agree (7). In addition, the impact of security concerns relevant to storing the payroll data outside the premises of the company was measured through the construct 'security concerns' captured by three items (Oliveira et al. 2014). Responses for this construct were measured in seven-point Likert-type scale with endpoints Not At All (1), Very Much (7). The impact of IT-Fashion was measured using two items relevant to the exposure of the firm towards fashion setters. Responses provided were measured using a seven-point Likert-type scale with endpoints Strongly Disagree (1), Strongly Agree (7).

Since Greek native speakers were targeted, the items were translated in the respondents' native language (Brislin 1970). The items translation was made by one of the authors of this paper and then reviewed and revised where necessary by the other two. The survey was pre-tested for clarity using MBA students who were familiar with the payroll business procedures and the cloud computing concepts. At a second phase, the survey was piloted by requesting six HR managers involved in the payroll procedure of their firm, to complete the survey while thinking aloud. After the end of the survey, items of the questionnaire at which respondents seemed to struggle were discussed, so as derive suggestions and reform the questions if necessary. At this stage, mini-interviews including open-ended questions were conducted with the pilot phase's participants, so as to ensure the validity of our understanding of the payroll procedure and its specificities.

4 Results

4.1 Data analysis and descriptive statistics

Various techniques and tools were used in order to test the research model and its hypotheses. Initially SPSS was used so as to organise the data and derive the descriptive analyses (illustrated in Table 1). To test the research model and the hypotheses proposed, Structural Equation Modelling (SEM) technique was employed using Partial Least Squares (PLS). PLS is considered appropriate if the sample size is small and the research aims at explaining the variance (Smith et al. 1997). As our sample size

number fulfilled the condition of being ten times the largest number of formative indicators used to measure a construct, it was considered safe and appropriate to use PLS for our analysis (Hair et al. 2011). To this end, Smart-PLS software was used with 1000 re-samples bootstrapping and exclusion of incomplete samples.

| Constructs | Descriptive Statistics | |
|--|------------------------|--------------------|
| | Mean | Standard Deviation |
| Reduction of costs (CST) | 4,27 | 1,54 |
| Remote access (RMA) | 3,97 | 2,11 |
| Security (SEC) | 4,56 | 2,13 |
| IT-Fashion (FSH) | 3,84 | 1,93 |
| Intention to adopt cloud-enabled payroll system (AD) | 1,56 | 0,92 |

Table 1. Descriptive statistics

4.2 Measurement Model Validation

Before proceeding with further examination of our research model, it is important that the measurement model's validity is tested. Measurements were addressed for their internal consistency, convergent validity and discriminant validity in accordance to testing systems suggested by existing literature (Fornell et al. 1981). All the constructs included in our model satisfied the criteria for internal consistency. This can be concluded as the Cronbach's alpha indicators for all of our constructs meet the .07 criterion (Nunnally et al. 1994). In addition, our model was successfully tested for convergent validity (that is testing that all items which measure a specific construct correlate). As suggested by Fornell and Lacker (1981) to meet a sufficient degree of convergent validity, it is essential that the average variance extracted (AVE) for each of our constructs is greater than .5. In addition, discriminant validity of the constructs was examined. This was achieved as the square roots of each factor's AVE were evaluated as greater than its correlations between it and the rest of other factors (Gefen et al. 2005). The analyses concerning the validity of our measurement model are described in Appendix B this paper. Since, our measurement model met the criteria discussed in this subsection, we proceeded with the structural validation stage of our analysis.

4.3 Structural Model Validation

Using PLS, the strength and significance of each hypothesis was tested. First of all, the path coefficients (pc) (or beta values, β) for each structural path were derived so as to understand the difference that each independent value makes in interpreting the variance in the depended variable. At the same time, the significance of each hypothesis was addressed by performing bootstrapping analysis on our dataset. In this way, the t-values (t) and corresponding significance for each relation were derived. Figure 2 illustrates our model in respect to the PLS values capturing path coefficients (pc), R^2 , t-values (t) and significances.

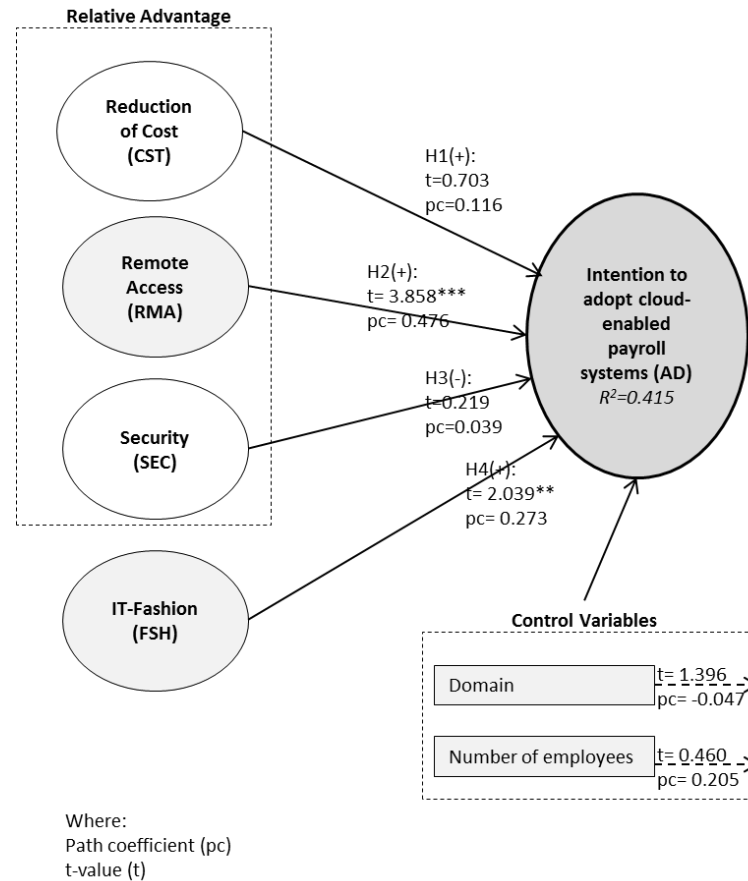


Figure 2. Structural model with results (** $p < .001$, ** $p < .05$.)

Squared multiple correlations, $R^2 = 0.415$, denote the good explanatory potential for the intention to adopt cloud-enabled payroll systems (AD) for the constructs used in our model (i.e. 41.5% can be explained by the constructs employed by our model). Two out of four of our hypotheses defined in Section 2 are supported at $<.05$ or even at $<.001$ level. More specifically, hypothesis 2 (H2), is strongly related to the intention to adopt cloud-enabled payroll systems as it has highly significant positive effects ($p < .001$) on the adoption intention. At the same time, hypothesis 4 (H4) is also supported as IT-Fashion seems to have a strong influence on the intention to adopt cloud-enabled payroll systems ($p < .05$). Regarding reduction of costs (H1) and security concerns (H3), our dataset did not reveal any significant influence on the adoption intention. Finally, the control variables employed for our analysis did not reveal any effect on the intention to adopt cloud-enabled payroll systems.

This section presented the measurement and model validation information concerning our study. The next section discusses the study results in more depth and outlines the implications arising by the hypotheses confirmed by the study.

5 Discussion

Our study introduced the concept of IT-Fashion and investigated its influence on the intention to adopt cloud computing and cloud-enabled payroll systems in particular. In addition, the study addressed the

relative advantage of cloud computing by analysing it into cost-savings, remote access and security concerns factors.

First, the strongest effect on the intention to adopt cloud services is the capability provided, through the use of cloud technology, to have remote access to the service from any place at any time. This factor is one of the capabilities that cloud computing provides which was not possible through previous technologies, hence it contributes to cloud's relative advantage. Although previous literature identifies this as an important factor influencing adoption (Janssen et al. 2011; Polyviou et al. 2014a), a comparison of its influence in relation to other relative advantage factors has not been studied. Using the case of cloud-enabled payroll systems, our study evaluated that remote access to the system is the most important factor influencing the intention to adopt (the significance of the finding was $p=.00003$). The preliminary interviews conducted for cloud services in general (Polyviou et al. 2014a), also converge to this finding. As denoted by the interviewees "use of information wherever you are with every device [...], wanted a CRM solution, something online so that accessible from everywhere" [quote of chief executive officer in Austria], "access from everywhere, any time [...], independence of physical site" [quote of chief information officer in Greece]. Hence, although in our study remote access is measured in the context of cloud-enabled payroll systems, we can conclude that it is an important factor grouped under cloud's relative advantage that is influencing cloud adoption more generally.

We also identified IT-Fashion to be impacting cloud adoption intentions. Organisations and its executives are exposed to fashion setter pressures. Hence, fashion setters are said to influence the decision makers of organisations as they impact the formulation of their perceptions towards the innovation. Hence, it could be anticipated that IT-Fashion has an influence on the decision to adopt the innovation. This is in line with the indications of existing literature on cloud computing, identifying that perceptions towards the term 'cloud' itself are also influencing the adoption intention (Morgan et al. 2013). This finding could not be clustered within the theoretical frameworks employed in previous studies of cloud computing even though authors called for further research on it (Morgan et al. 2013). By employing cloud-enabled payroll systems as our example, throughout our study we have demonstrated that the tendency to follow IT-Fashion has a remarkable influence on the intention to adopt. This is confirmed also by our preliminary research study that targeted the adoption of cloud services in general (Polyviou et al. 2014a). For example, some of our interviewees noted that the cloud product was spotted through a "CRM solutions fair" [quote of chief information officer in Austria], "the need was [...] resolved in a more traditional/pen and paper manner [...but then saw], it on facebook, watched the youtube video to understand how it works"[quote of chief executive officer in Greece]. Thus, we can conclude that IT-Fashion influences cloud adoption.

In addition, under relative advantage, reduction of cost was another factor hypothesised to be influencing the intention to adopt cloud computing. As cloud computing minimises investment on IT hardware, IT support cost and reduces capital expenditure, it was expected to be positively influencing the intention to adopt. This was also indicated by the findings of Oliveira et al. (2014). However, according to the results of our study cost reduction is not influential if compared to the influence of remote access and IT-Fashion factors. One possible justification of this finding is the fact that the needs of a firm in terms of payroll system licences are standard for each year, unlikely to vary across time and need limited computation power. Due to this characteristic of payroll systems we call for further investigation to address if there is diversity of answers when considering systems of a less predicted pattern of licence needs.

Moreover, given that firms are often handling sensitive data (e.g. financial statements, employee personal information etc), it was hypothesised that security concerns negatively influence cloud adoption. Using the case of cloud-enabled payroll systems, our study did not reveal that security concerns have an influence on the intention to adopt. Given that payroll systems are handling some of the most sensitive data of the organisation; this finding could be an indication that organisations and service providers are starting to build a more trusted relationship.

Also, previous research studies attempt to diversify their results based on the sector to which responding organisation belongs to (Oliveira et al. 2014). Hence, our study included this attribute as the control variable ‘domain’ and captured respondents from more than twelve different domains. Our research results concluded that the domain of the organisation did not have an influence on the intention to adopt cloud and cloud-enabled payroll systems in particular. Furthermore, the existence of the HR department in the organisation was used as a control variable, as well as the number of employees. We consider the number of employees a variable measuring the size of the organisation. We opted against the use of revenue as a control variable, as the current economic situation in Greece has reduced the validity of this measure as a proxy for organizational size.

Overall, our study demonstrated the relevance of IT-Fashion for cloud adoption as IT-Fashion hypothesis of our model explains a significant percentage of the adoption intentions. Although our research model was tested in cloud-enabled payroll systems, supportive insights from our preliminary qualitative studies targeting the adoption of cloud services in general illustrate the applicability of our findings for cloud adoption literature more generally. Although IT-Fashion explains a remarkable proportion of the intention to adopt cloud, it has been overlooked by previous efforts that endeavour to explain the factors influencing the intention to adopt cloud services. Our study has contributed by shedding further light on cloud adoption decisions and manifesting the relevance of IT-Fashion as another key factor that influences cloud adoption intention. In addition, the study specified the relative advantage of cloud by identifying factors that define cloud’s relative advantage and testing their relevance. In this way, it has been clarified that remote access is the most important factor of relative advantage influencing the intention to adopt cloud, at least in the case of payroll systems.

Beyond its theoretical contribution, the study generated implications that can be valuable to practitioners especially from a marketing perspective. First of all, the relevance of IT-Fashion confirms the need for cloud vendors to invest in marketing campaigns. Such marketing campaigns could aim at creating a positive perception towards cloud. As indicated by our findings, the views of decision makers are highly influenced by fashion setters. Hence, targeting potential customers through fairs, press releases and social media can contribute to raising the vendor’s market share. In addition, our findings provide indications about the content of marketing campaigns. More specifically, according to the insights of our study, marketing campaigns should focus on promoting certain benefits of the service that are enabled by cloud. For example, remote access was identified as the most important factor influencing the adoption intention, at least for cloud-enabled payroll systems. Hence, vendors offering cloud-enabled payroll systems should consider the promotion of the remote access functionality of their products compared to other characteristics such as its cost-effectiveness. Finally, heavy investments in providing security assurances seem to be no longer necessary for vendors, at least for those offering payroll systems.

We acknowledge that our study has a number of limitations. First, the sample targeted Greek companies, which implies that the study might reflect a situation in a specific country. Furthermore, although the findings of our study were also supported by the findings of preliminary study which targeted cloud service in general, the core conclusions of the study were demonstrated by using the example of cloud-enabled payroll systems. In addition, the model aimed at examining if IT-Fashion influences the intention to adopt cloud. This was achieved by comparing it to the factors that have been found in the extant literature to have the highest impact on the intention to adopt (i.e. the technology’s relative advantage). As a result, the model did not capture additional factors that may fall out of cloud’s relative advantage, but may have a smaller influence on the adoption intention.

6 Conclusion

This paper has studied the impact of IT-Fashion on the cloud adoption intention. To achieve this, the study has followed the strategy of juxtaposing it to cloud’s relative advantage. Future research should consider enhancing the existing research model with additional factors, in addition to those grouped under cloud’s relative advantage. By juxtaposing additional factors to IT-Fashion, the influence of IT-

Fashion could be further compared and factors influencing the intention to adopt cloud can be prioritised. In the study presented in this paper, by employing cloud-enabled payroll systems as our example, we have demonstrated that IT-Fashion influences the intention to adopt cloud to a remarkable extent. In addition, we have analysed cloud's relative advantage into more specific factors, revealing that remote access is the cloud characteristic that makes the technology more advantageous compared to what it attempts to supersede. We have discussed our results in more depth in the light of earlier research on cloud computing adoption. To confirm the generalizability of our results for cloud computing adoption, future research could focus on examining the impact of IT-Fashion using other types of cloud services, preferably of a more strategic importance. Finally, analysing the IT-Fashion variable into further sub-variables could provide additional insights on the most effective marketing strategies that cloud providers can follow.

Acknowledgments

This research has been co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Thalis - Athens University of Economics and Business - Software Engineering Research Platform. The authors also acknowledge the feedback of Dr. Stephanos Androutsellis-Theotokis on earlier drafts of this work.

Appendix A

| | | | |
|----------------------|------|--|-------------------------|
| Reduction of Cost | CST1 | The cost of adoption of a cloud-enabled payroll system are far greater than the benefits. | (Premkumar et al. 1999) |
| | CST2 | The cost of maintenance and support of a cloud-enabled payroll system are very high for our business. | (Premkumar et al. 1999) |
| | CST3 | The amount of money and time invested in training employees to use a cloud-enabled payroll system are very high. | (Premkumar et al. 1999) |
| Remote Access | RMA1 | My firm would be interested in having remote access to the payroll system. | (self-developed) |
| | RMA2 | The remote access functionality of the payroll system will improve the effectiveness of the tasks completed by the HR department. | (self-developed) |
| | RMA3 | The remote access functionality of the payroll system will raise the productiveness of the HR department | (self-developed) |
| Security Concerns | SEC1 | Degree of concern of my firm for the security of the payroll data in the cloud | (Oliveira et al. 2014) |
| | SEC2 | Degree of concern of the employees of my firm for the security of payroll data in the cloud | (Oliveira et al. 2014) |
| | SEC3 | Degree of concern of employees of my firm about the privacy of the payroll data in the cloud | (Oliveira et al. 2014) |
| IT-Fashion Influence | FSH1 | We were one of the first ones to find out about cloud-enabled payroll systems, since we are informed about the technological applications of our domain through conferences, press releases and social media | (self-developed) |
| | FSH2 | We have heard about cloud-enabled payroll systems since a long time ago, as we are frequently following the technological applications of our domain through conferences, press releases and social media | (self-developed) |
| Intention to Adopt | AD1 | At what stage of adoption is your firm currently engaged regarding cloud-enabled payroll systems? (Not considering; Has evaluated, and does not plan to adopt; Currently evaluating (e.g., in a pilot study); Has evaluated and plans to adopt; Has already adopted such systems) | (Thiesse et al. 2011) |
| | AD2 | If you are expecting that your firm will adopt cloud payroll systems in the future, when do you think it will happen? (Not considering to adopt; In more than 5 years; Within 2-5 years; Within 1-2 years; In less than a year) | (Thiesse et al. 2011) |

Table A . List of Constructs and Items

Appendix B

| | Cronbach's Alpha | Composite Reliability | AVE | CST | RMA | SEC | FSH | AD |
|--|------------------|-----------------------|-------|--------|--------|--------|-------|-------|
| Reduction of costs (CST) | 0.786 | 0.829 | 0.621 | 0.788 | | | | |
| Remote access (RMA) | 0.917 | 0.948 | 0.858 | 0.098 | 0.926 | | | |
| Security (SEC) | 0.934 | 0.958 | 0.884 | -0.072 | 0.180 | 0.940 | | |
| IT-Fashion (FSH) | 0.911 | 0.957 | 0.917 | 0.147 | -0.358 | -0.220 | 0.957 | |
| Intention to adopt cloud-enabled payroll system (AD) | 0.738 | 0.880 | 0.787 | 0.152 | 0.554 | -0.262 | 0.268 | 0.887 |

Table B. Convergent Validity

References

- Abrahamson, E. 1991. "Managerial Fads and Fashions: The Diffusion and Rejection of Innovations," *The Academy of Management Review* (16:3), pp 586-612.
- Abrahamson, E. 1996. "Management Fashion," *The Academy of Management Review* (21:1), pp 254-285.
- Alshamaila, Y., Papagiannidis, S., and Li, F. 2013. "Cloud computing adoption by SMEs in the north east of England," *Journal of Enterprise Information Management* (26:3), pp.250-275.
- Anandasivam, A., Best, P., and See, S. 2010. "Customers' Preferences for Infrastructure Cloud Services," *In Proceedings of the 12th IEEE Conference on Commerce and Enterprise Computing, Shanghai, China*.
- Bohn, R. B., Messina, J., Liu, F., Tong, J., and Mao, J. 2011. "NIST Cloud Computing Reference Architecture," *In Proceedings of the 2011 IEEE World Congress on Services*, IEEE Computer Society, pp. 594-596.
- Brislin, R. W. 1970. "Back-Translation for Cross-Cultural Research," *Journal of Cross-Cultural Psychology* (1:2), pp 185-216.
- Carr, N. 2009. *The Big Switch: Rewiring the World, from Edison to Google*, (1 ed.) W.W. Norton Company London, U.K.
- Chau, P. Y. K., and Tam, K. Y. 1997. "Factors Affecting the Adoption of Open Systems: An Exploratory Study," *MIS Quarterly* (21:1), Zmud, B. (ed.), pp 1-24.
- Durkee, D. 2010. "Why Cloud Computing Will Never Be Free," *Communications of the ACM* (53:5), pp 62-69.
- Fornell, C., and Lacker, D. F. 1981. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research* (18:1), pp 39-50.
- Gefen, D., and Straub, D. 2005. "A Practical Guide To Factorial Validity Using PLS-Graph: Tutorial And Annotated Example," *Communications of the Association for Information Systems* (16:1), pp 91-109.
- Giessmann, A., and Stanoevska, K. 2012. "Platform As A Service – A Conjoint Study On Consumer's Preferences," in *In Proceedings of the 32th International Conference of Information Systems(ICIS), Orlando, U.S.A.*
- Hair, J. F., Ringle, C. M., and Sarstedt, M. 2011. "PLS-SEM: Indeed a Silver Bullet . *Journal of Marketing Theory and Practice* (19:2), pp 139-152.
- Janssen, M., and Joha, A. 2011. "Challenges for Adopting Cloud-based Software as a Service (SaaS) in the Public Sector," in *19th European Conference on Information Systems (ECIS), Helsinki, Finland*.
- Joint, A., Baker, E., and Eccles, E. 2009. "Hey, you, get off of that cloud?," in *Computer Law & Security Review*, pp. 270-274.
- Kieser, A. 1997. "Rhetoric and Myth in Management Fashion," *Organization* (4:1), pp 49-74.
- Köhler, P., Anandasivam, A., and Ma, D. Year. "Cloud Services from a Consumer Perspective," *16th Americas Conference on Information Systems (AMCIS), Lima, Peru, 2010*.
- Kuan, K. K. Y., and Chau, P. Y. K. 2001. "A perception-based model for edi adoption in small businesses using a technology-organization-environment framework," *Information & Management* (38:8), pp 507-521.
- Low, C., and Chen, Y. 2011. "Understanding the determinants of cloud computing adoption," *Industrial Management & Data Systems* (111:7), pp 1006-1023.
- Morgan, L., and Conboy, K. Year. "Factors Affecting The Adoption Of Cloud Computing: An Exploratory Study," *21st European Conference on Information Systems (ECIS), Utrecht The Netherlands*.
- Nunnally, J. C., and Bernstein, I. H. 1994. *Psychometric Theory*, McGraw-Hill: New York.

- Oliveira, T., Thomas, M., and Espadanal, M. 2014. "Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors," *Information & Management* (51:5), pp 497-510.
- Polyviou, A., Pouloudi, N., and Pramadari, K. 2014a. "Cloud Adoption: Relative Advantage or IT fashion?," *European Conference on Information Systems (ECIS), Tel-Aviv, Israel*.
- Polyviou, A., Pouloudi, N., and Rizou, S. 2014b. "Which Factors Affect Software-as-a-Service Selection the Most? A Study from the Customer's and the Vendor's Perspective," in *In Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS), Hawaii, U.S.A.*
- Premkumar, G., and Roberts, M. 1999. "Adoption of new information technologies in rural small businesses," *Omega - The International Journal of Management Science* (27:4) 8//, pp 467-484.
- Rogers, E. M. 1995. *Diffusion of Innovations*, Free Press: New York.
- Smith, J. B., and Barclay, D. W. 1997. "The effects of organizational differences and trust on the effectiveness of selling partner relationships," *Journal of Marketing* (61:1), pp 3-21.
- Swanson, E. B. 2012. "The Manager's Guide to IT Innovation Waves," *MIT Sloan Management Review* (53:1), pp 75-83.
- Swanson, E. B., and Ramiller, N. C. 1997. "The Organizing Vision in Information Systems Innovation," *Organization Science* (8:5), pp 458-474.
- Thiesse, F., Staake, T., Schmitt, P., and Fleisch, E. 2011. "The rise of the "next-generation bar code": an international RFID adoption study," *Supply Chain Management: An International Journal* (16:5), pp 328-345.
- Tornatzky, L. G., and Fleischer, M. 1990. *The processes of technological innovation*, Lexington Books: Massachusetts.
- Wang, P. 2010. "Chasing the Hottest IT: Effects of Information Technology Fashion on Organizations," *MIS Quarterly* (34:1), pp 63-85.
- Whitley, E. A., Mooney, J., Retana, G., Reynolds, P., and Venters, W. 2013. "Panel 5: the shift to cloud computing: a slow train comin'?", in *In Proceedings of the International Conference on Information Systems (ICIS), Milan, Italy*.
- Willcocks, L., Venters, W., and Whitley, E. 2014. *Moving to the Cloud Corporation*, (1st ed.) Palgrave Macmillan: Hampshire, U.K.
- Yusof, M. M., Kuljis, J., Papazafeiropoulou, A., and Stergioulas, L. K. 2007. "An evaluation framework for Health Information Systems: human, organization and technology-fit factors (HOT-fit)." *International Journal of Medical Informatics* (77:6), pp 386-398.

CRYPTOCURRENCY VALUE FORMATION: AN EMPIRICAL ANALYSIS LEADING TO A COST OF PRODUCTION MODEL FOR VALUING BITCOIN

Hayes, Adam, The New School for Social Research, New York NY, hayea411@newschool.edu

Abstract. *This paper aims to identify the likely source(s) of value that cryptocurrencies exhibit in the marketplace using cross sectional empirical data examining 66 of the most used such 'coins'. A regression model was estimated that points to three main drivers of cryptocurrency value: the difficulty in 'mining' for coins; the rate of unit production; and the cryptographic algorithm employed. These amount to relative differences in the cost of production of one coin over another at the margin, holding all else equal. Bitcoin-denominated relative prices were used, avoiding much of the price volatility associated with the dollar exchange rate. The resulting regression model can be used to better understand the drivers of relative value observed in the emergent area of cryptocurrencies. Using the above analysis, a cost of production model is proposed for valuing bitcoin, where the primary input is electricity. This theoretical model produces useful results for both an individual producer, by setting breakeven points to start and stop production, and for the bitcoin exchange rate on a macro level. Bitcoin production seems to resemble a competitive commodity market; in theory miners will produce until their marginal costs equal their marginal product.*

Keywords: Bitcoin, cryptocurrencies, altcoins, asset pricing.

1 Empirical Analysis to Define Which Factors are Determinants in Cryptocurrency Value Formation

1.1 Introduction

Due to Bitcoin's growing popular appeal and merchant acceptance, it has become increasingly important to try to understand the factors that influence its value formation. However, price fluctuations of bitcoin versus national currencies such as the U.S. dollar, euro or Chinese yuan, have been extremely volatile. This extreme price volatility produces a lot of noise which makes meaningful analysis difficult. In fact, there is increasing evidence that the rise in price for one bitcoin to over \$1,000 around December 2013 was largely caused by coordinated price manipulation at the Mt. Gox exchange involving fraudulent trading algorithms which pilfered customer accounts.¹ The subsequent failure of the Mt. Gox exchange and the associated customer accounts was likely a direct result of this market manipulation. Fortunately, there is an active and fairly liquid market for various altcoin-bitcoin trading pairs. By looking at bitcoin-denominated relative prices and removing the external dollar, euro, yuan, etc. exchange rates, much of the noise and price volatility can be removed, making for a much better analysis of the data. Comparing how the variations in several shared attributes of cryptocurrencies affects their relative prices with bitcoin, factors that influence value formation can be identified.

This paper describes a cross-sectional data analysis of 66 cryptocurrencies in such a manner using objective factors shared by each one of them. The findings indicate that relative value formation occurs in production at the margin, much like other commodities.

¹ See: *The Willy Report*: <https://willyreport.wordpress.com/>

1.2 A brief overview of Bitcoin

The technical specifications of the Bitcoin and altcoin protocols are beyond the scope of this paper, however some key points must be understood before going any further, under the assumption that many readers have little to no prior knowledge of this topic.

Taking bitcoin as the generic example, one can then extend those concepts to the greater universe of altcoins. This overview is purposefully brief and meant only to clarify some points that will be referred to in this paper.

Bitcoin is an open source software-based online payment system that emerged in 2008-2009. Payments are recorded in a shared public ledger using its own unit of account, which is also called bitcoin, symbolically BTC.²

Transactions occur peer-to-peer without a central repository or single administrator – it is a decentralized virtual currency which also can be completely anonymous. New bitcoins are created as a reward for transaction processing work in which users offer their computing power to verify and record payments into the public ledger. Also known as “mining”, individuals or companies engage in this activity in exchange for the chance to earn newly created blocks of bitcoins.

Mining is carried out by specialized hardware which has a certain amount of computational power, measured in *hashes* per second.³ The aggregate bitcoin network has a cumulative computational power additive of all the mining effort employed around the world. For every one GigaHash/second (GH/s) any individual miner puts online, for example, that amount will be added to the overall network power.⁴ Mining is quite competitive, in the sense that somebody mining with more computational power or with greater efficiency has a better chance of finding a block than somebody with less. Computational effort in cryptocurrency production is often referred to as alternatively *hashpower*, *hashing power*, *mining effort*, or *hashrate*.

Besides mining, bitcoins can be obtained in exchange for currencies such as dollars, euros, etc., for other altcoins, or in exchange for products, and for services. Users can send and receive bitcoins electronically using 'wallet' software on a personal computer, mobile device, or a web application.

1.3 A short survey of relevant literature

There is a new and emerging academic literature regarding cryptocurrencies, with most emphasis surrounding Bitcoin. Much of the economic study undertaken has attempted to address the “moneyness” of bitcoin or whether it is more analogous to a fiat versus commodity money, like a 'digital gold' (Gertchev, 2013) (Harwick, 2014) (Bergstra, 2014).

Yermack (2013) looks at bitcoin's moneyness and points out weaknesses in bitcoin as a currency. Yermack claims that bitcoin (and all cryptocurrencies by association) have no intrinsic value. I consider the potential that while its characteristics are intangible and the labor employed to mine for them is computational rather than human or mechanical, a bitcoin does indeed have an intrinsic value, albeit virtual, which cannot be directly compared to tangible intrinsic value possessed by gold, for example. I don't disagree with the premise that bitcoin and its cousins are not money in the strict sense and that many issues stand in the way of it moving

² Bitcoin with a capital “B” refers to the protocol, network and system; while bitcoin with a small “b” refers to individual units of the cryptocurrency.

³ Hashes are somewhat analogous to the processing power of a CPU microchip, which is measured in hertz in defining how many individual computations can be achieved per second.

⁴ A GigaHash/second (GH/s) is one billion hashes per second.

toward mass acceptance and appeal. Yermack makes a very valid point that the price volatility of Bitcoin as expressed in dollars is quite high and that its dollar price may vary significantly among the various exchanges. He mentions that this can cause problems when trying to analyze price data.

For this paper, I have used only bitcoin as the denomination for the various cryptocurrency prices, without the need for dollars. Of course, one can then transpose all prices to dollars using a current dollar-bitcoin exchange rate if they chose. Hence, bitcoin is always worth 1 BTC, and all other cryptocurrencies expressed in decimal form as $x.xxxxxxxx$ BTC. It is worth noting that for many of these cryptocurrencies there only exists pairwise trading on exchanges between itself and BTC (or another cryptocurrency); there are far less altcoin/USD trading pairs than altcoin/BTC pairs. Attempts thus far at valuation, or sources of value, have focused almost entirely on bitcoin without consideration to the scope of alternative cryptocurrencies or altcoins.

Hanley (2013) argues that the value of bitcoin floats against other currencies as a pure market valuation with no fundamental value to support it. Woo, et al. (2013) proposes that bitcoin may have some fair value due to its money-like properties as a medium of exchange and a store of value, but without any other underlying basis.

Jenssen (2014) identifies the “proof-of-work” feature of the mining protocol, implying there may be some sort of computer-labor power source of value. Jenssen also argues that the observed market price of bitcoin in dollars is due to demand given a limited supply. The fact that there will only be 21 million bitcoins as a bounded limit on eventual supply could very well be a red herring; since each bitcoin is divisible to eight decimal places and that number of decimal places can be theoretically increased. There is nothing to prevent the functional unit from being a nano-bitcoin, for example. Although dealing with leading zeros might be cumbersome, it is not prohibitive. With traditional money, there is no effective way to have the functional unit as a fraction of a cent. This paper shows that what is more important as a source of value seems to be the rate of unit formation.

Van Alstyne (2014) considers a source of bitcoin value to be the technological value in solving the so-called double spend problem. While this breakthrough has certainly allowed for the viability of bitcoin, it does not in and of itself make for value. For why then would other cryptocurrencies, which have the same or similar protocols underlying them, have disparate relative values?

Bouoiyour & Selmi (2014) attempt to describe bitcoin value by regressing its market price against a number of independent variables including those such as the market price of gold, occurrences of the word 'bitcoin' in Google searches, the velocity of bitcoin measured by transaction data, and so on. Largely, the variables when regressed were not statistically significant at the 5% or better level of significance. Lags on the price of Bitcoin itself were found to carry some weight, but that can be an artefact of the time-series analysis. Seemingly, only the regression on lagged Google search results were significant at the 1% level. While this finding is interesting, it shows that many variables which may be hypothesized to confer value actually do not. In fact, in an 18-variable multiple regression the R^2 value they obtained was only 0.4586, indicating that some other variables must account for over half of bitcoin's dollar value. Because cryptocurrencies are nascent and still highly speculative and volatile, using time series analysis can be misleading and uninformative over the short life time of its existence.

Polasik et al. (2014) concludes that bitcoin price formation is the result primarily of its popularity and the transactional needs of its users. They, too, utilized Google search results and found this variable to be highly significant, while the number of transactions (a proxy for velocity) was found not to be. I argue that use of Google search results is not a good metric and that the found correlation might be spurious. In the period when these studies took place, the dollar price of bitcoin was rising rapidly. This rapid price increase caused increasing media attention and word-of-mouth introducing it to more and more people who subsequently searched the internet to gain more information. The people actively mining for or transacting in bitcoin, I surmise, would not need to repeatedly input the word 'bitcoin' as a Google search

term, rather people looking at it for the first time, or to investigate it to a greater degree would utilize such a search.

Zhang et al (2014) looks at alternative cryptocurrencies (altcoins) in conjunction with bitcoin, however they only consider three such altcoins (litecoin, dogecoin and reddcoin). Their work is largely descriptive, but lays the groundwork for future research on cryptocurrencies in general and in the framework of micro- and macroeconomics.

Gandal and Halaburda (2014) analyze the competition among a small number of cryptocurrencies in the marketplace and competition between four online exchanges. They found that arbitrage opportunities, for the most part, do not exist. The small sample size makes their findings a bit incomplete; they also relate cryptocurrency prices to the dollar instead of using bitcoin as the base for comparison. Due to a number of frictions in transactions between cryptocurrencies and national fiat money, markets tend to be more efficient and less volatile when looking at cryptocurrencies relative to a bitcoin base.⁵ This transactional friction and the noise it creates may also be why it was found that gross trading opportunities were much greater across exchanges than within exchanges – where conversions to and from fiat currencies are required.

Garcia, et al. (2014) asserts that the cost of production through mining does matter in coming up with a fundamental value for bitcoins insofar as it represents a lower bound. This paper will elaborate on that general idea and formalize it to identify a cost of production model for bitcoin. Doing so can identify theoretical break-even levels in market price, electricity cost, mining energy efficiency, and mining difficulty for individual miners – and may be extended to impute averages for the aggregate network.

While it may be tempting to objectify these results to impute a true intrinsic value for bitcoin, I would caution against making such a leap. Even if the models developed in this paper can theoretically determine an intrinsic value, extreme volatility and frequent market price fluctuations in the few years since bitcoin has been around could make identifying such an intrinsic value meaningless in application. There is also the matter of subjective components of value formation which are more difficult to quantify.

1.4 Assumptions and hypotheses

I will use Bitcoin as the generic example to explain the more general case of cryptocurrencies. There are a few fundamental variables that have been hard-wired into the Bitcoin protocol at its inception. As most altcoins share a common Bitcoin lineage, the majority of cryptocurrencies have the same set of built-in variables. The numerical values of these variables can be thought of as arbitrary to some extent when they were created. These variables include:

1- The total number of “coins” ever to be created. For bitcoin, this value will be 21,000,000 and no more. I will refer to this variable as *Total Money Supply*.

2- Each block found by mining will contain a specified number of units. A block of bitcoins initially contained 50 BTC, currently it stands at 25 BTC per block, and that amount will continue to be halved over time, approximately every four years. I will call this variable representing the number of coins in a block the *Block Reward*.

⁵ Transaction costs & fees, regulatory issues, time waiting for bitcoin confirmations, and time waiting to clear fiat money deposits/withdrawals are just some of these frictions

3- A block of coins will be found by mining over the same interval, on average, regardless of the magnitude of mining effort. Bitcoin blocks will be found, on average, once every 10 minutes. I will refer to this variable as *Block Time*.

4- The network will check to ensure that the specified Block Time as been achieved on average over some number of blocks previously mined. In the case of bitcoin, after 2,016 blocks have been found, the system will check and see if the actual average time in creating blocks was greater or less than 10 minutes. If it was less than 10 minutes, the system will increase the marginal difficulty in finding new blocks so that the 10 minute average will be restored. This I will call the *Difficulty Retarget*.

5 – The underlying *Algorithm* is the cryptologic hash function used as the basis for the protocol. Bitcoin uses what is known as SHA-256d.⁶ Many altcoins use that method, while others use a function called scrypt. The inner workings of the algorithms used are beyond the scope of this paper.

6 – The *Difficulty* variable is exogenous and describes how hard (in computational power) it is to find a new block given a fixed level of hashpower. Because of the Difficulty Retarget mechanism, the difficulty will adjust up or down as aggregate mining effort is employed or removed from the network.

7 – The market *Price* is the observable price on exchanges where altcoin/BTC trading pairs are listed.

By endowing a cryptocurrency with a steady and known rate of unit formation, it cannot be influenced by any central authority. It is important to note that by employing more computational power (e.g. mining hardware) to the network, it may temporarily increase the likelihood that the individual miner with the most power will be most productive; however, the network will check the Difficulty Retarget and adjust the Difficulty accordingly to restore the Block Time. Therefore, if hypothetically somebody were to put online the most powerful new technology, say many Peta-Hashes/second (1,000,000s GH/s) of computational power, once the network detects that the average time between block creation was too low it would adjust the difficulty up accordingly, rendering that new technology merely adequate, and also rendering every other miner's technology inferior or even obsolete.

In devising new and alternative cryptocurrencies, the creator of a fresh 'coin' need only look at the open source computer code, copy it, and change one or more of the above variables to suit their liking. Thus, there are is a diverse universe of altcoins: some that have only a 1 Difficulty Retarget instead of 2,016; some which set the Total Money Supply to either a small handful, or any number including an infinite amount; some set the Block Reward to a fraction of a coin per block while others issue many thousands of coins per block; virtually any combination conceivable.

Because there are active markets on the internet, exchange ratios and prices for each of these altcoins is known and are tradeable in real-time and across a number of platforms. The open source nature of the underlying code also makes finding the values for the above variables easy to obtain.

The fact that there are altcoins with all sorts of configurations makes it a rich data set with which to inquire into what factors may be determinants of value on to them.

⁶ SHA-256d and scrypt remain the most commonly used mining algorithms. New algorithms such as X11 exist too but for this study only SHA and scrypt coin data are used for simplicity.

a priori, my hypotheses are:

H1. The amount of mining (computational) power devoted to finding a 'coin' is positively correlated to altcoin value. The more aggregate computational power employed in mining for a cryptocurrency, the higher the value. I make this assertion for a number of reasons. First, the more mining power there is, the more acceptance for that 'coin' can be inferred – since mining also serves to verify transactions, the amount of mining power in use is a proxy for overall use and acceptance of that altcoin.

A cryptocurrency with no acceptance or usage will have neither value nor computational power directed at it. Second, a rational miner, motivated by profit, would only seek to employ mining resources to a profitable pursuit. Therefore, if the marginal cost of mining exceeded the marginal price of mining, that miner would redeploy his resources elsewhere, removing the computational power from the network of that altcoin and into another. Third, the computational power is a proxy for the mining difficulty since the more network power employed, the greater the difficulty will become in order to maintain the pre-programmed Block Time. Therefore, difficulty can be used as an indirect proxy of aggregate mining power.

There is the possibility that the causal relationship between price and computational power is reversed, or bidirectional. It is certainly plausible that computational power will be deployed to where it is already profitable to do so (e.g. prices are already high). To check this, a Granger causality test was run on price and aggregate hashpower. The results strongly indicate that causality runs one-way from mining effort to price and not the other way.

H2. The rate of 'coins' found per minute is negatively correlated to altcoin value. Extending the law of diminishing marginal utility, the more readily something is available, and the more rapid that pace of availability, the lower the value; in other words, the faster the rate of unit formation, the lower the price. If an altcoin is configured such that it produces an abundance of units per block, and/or blocks are found in rapid succession, it will negatively impact the value of those units. On the other hand, scarcity per block would tend to lead to greater perceived value. This hypothesis takes into account the variables of Block Reward and Block Time.

H3. The percentage of coins mined thus far compared to that which is left to be mined before the Total Money Supply is reached is positively correlated to altcoin value. Since there is an exogenous future limit to the money supply, the closer the percentage of units that have been mined compared to what is still left to be found will increase its scarcity and confer value. This can be computed by dividing the number of coins found So far to date by Total Money Supply. This can be used to measure relative scarcity.

H4. Altcoins based on the scrypt algorithm will be more valuable than SHA-256d, all else equal. The scrypt system was put into use with cryptocurrencies in an effort to improve upon the SHA-256d protocol which preceded it and which bitcoin is based on. Specifically, scrypt was employed as a solution to prevent specialized hardware from brute-force efforts to out-mine others for bitcoins. As a result, scrypt altcoins require more computing effort per unit, on average, than the equivalent coin using SHA-256d. The relative hardness of the algorithm confers relative value.

H5. The longevity of the cryptocurrency is positively related to altcoin value. In other words, the longer a cryptocurrency has been around and used, the more value it will have. This is because in a competitive environment, such as that in altcoins, the 'losers' will simply cease to exist. Therefore, the longer a cryptocurrency has persisted, the more valuable it should be. All cryptocurrencies have a 'genesis' date which is easy to ascertain.

1.5 Empirical results of regression analysis

A least-squares (OLS) multiple regression was estimated using cross-sectional data from 66 of the most widely used and actively traded altcoins with the following specification:

$$\ln(\text{PRICE}) = \beta_1 + \beta_2 \ln(\text{GH/s}) + \beta_3 \ln(\text{COINS_PER_MIN}) + \beta_4 (\% \text{COINS_MINED}) + \beta_5 (\text{ALGO}) + \beta_6 (\text{DAYS_SINCE}) + e$$

where:

$\ln(\text{PRICE})$ is the natural logarithm of the bitcoin-denominated market price on September 18, 2014.

$\ln(\text{GH/s})$ is the natural logarithm of the computational power in GigaHashes per second.

$\ln(\text{COINS_PER_MIN})$ is the natural logarithm of the number of coins found per minute, on average which is computed by dividing Block Reward and Time Between Blocks.

$\% \text{COINS_MINED}$ is the percentage of coins that have been mined thus far compared to the total that can ever be found.

ALGO is a dummy variable for which algorithm is employed, taking on the value of '0' if SHA-256 and '1' if scrypt.

DAYS_SINCE is the number of calendar days from inception of the cryptocurrency through September 18, 2014.

The resulting regression output produced *Model A*:

$$\ln(\text{PRICE}) = -9.68^{***} + 0.67 \cdot \ln(\text{GH/s})^{**} - 0.98 \cdot \ln(\text{COINS_PER_MIN})^{***} - 0.57 \cdot \text{COINS_MINED} + 7.43 \cdot \text{ALGO}^{***} + 0.00067 \cdot \text{DAYS_SINCE}$$

$R^2 = 0.844$, Adjusted $R^2 = 0.830$, DW-statistic = 2.24, F-statistic = 63.71

t-statistics are indicated according to each explanatory variable. *** indicates $p < 0.001$, ** indicates $p < 0.005$

The R^2 is quite high, suggesting that approximately 84.4% of the variation in relative cryptocurrency prices are determined by the variables in the model.

Hypothesis *H1* is supported in that the coefficient is positive as expected *a priori* (prices increase as computational power increases), and the t-statistic indicates that it is highly statistically significant that computational power influences price.

Hypothesis *H2* is supported in that the coefficient is negative as expected *a priori* (prices decrease as the rate of coin production per minute increases), and the t-statistic indicates that it is highly statistically significant that coins produced per minute influences price.

Hypothesis *H3* is *not* supported in that the sign of the coefficient is unexpected, and also the t-statistic indicates that percentage of coins mined is not statistically significant. One possible reason for this result is that while the total number of coins is determined at the inception of a cryptocurrency, the 'coins' themselves are divisible down to 8 decimal places by default, and that number of decimal places can be increased, potentially without limit. Therefore, it may be the case that an absolute Total Money Supply may not actually be a limiting factor since once that ceiling is reached, the units can simply be divided and subdivided. For example, 1 BTC is actually 1.00000000 BTC, and there is nothing preventing 0.00000001 BTC from having useful value (except perhaps that it is cumbersome).

Hypothesis *H4* is supported in that the coefficient is positive as expected *a priori* that scrypt altcoins are more valuable than SHA-256, on average, and the t-statistic indicates that it is highly statistically significant that scrypt as opposed to SHA-256 influences price.

Hypothesis *H5* is *not* supported by the regression output, although the sign of the coefficient is positive which was expected *a priori*, the number of days since inception is not statistically significant. One possible reason for this result is that the vast majority of altcoins are less than two years old, which hasn't given the market enough time for competition to weed out the losers and reward the winners.

Removing the independent variables that were not statistically significant in *Model A*, a new regression was estimated to produce *Model B*, which had the following output:

$$\ln(\text{PRICE}) = -9.53^{***} + 0.69 \cdot \ln(\text{GH/s})^{***} - 0.98 \cdot \ln(\text{COINS_PER_MIN})^{***} + 7.46 \cdot (\text{ALGO})^{***}$$

$R^2 = 0.843$, Adjusted $R^2 = 0.835$, DW-statistic = 2.12, F-statistic = 111.04

t-statistics according to each explanatory variable and full regression outputs available in the appendix.

*** indicates $p < 0.001$.

Model B represents a more parsimonious output with a very similar R^2 compared to *Model A*, while improving the F-statistic and slightly improving the t-statistics for each explanatory variable. The model was checked for consistency with the assumptions of a linear regression, and exhibits normality of residuals, does not exhibit heteroscedasticity, collinearity, or other common regression errors.

Model B infers that holding all else constant:

- given a 1% increase in aggregate GH/s output, the price will rise by approximately 0.69%.
- given a 1% increase in coins produced per minute, the price will fall by approximately 0.98%.
- given that the altcoin uses the scrypt protocol, the price will be higher by approximately 7.46% compared to its SHA-256 counterpart., all else equal.

I would argue that in either of these regression models the intercept term has no valid economic interpretation.

1.6 Discussion of results

These econometric models can be useful in a number of ways. It specifies the factors that influence relative prices across a wide variety of cryptocurrencies that exist, inclusive of Bitcoin, and without the noise generated by price volatility with exchange rates against national currencies. Using these findings, pricing existing or newly created cryptocurrencies can be undertaken with some greater degree of confidence.

It shows that more than 84% of relative value formation can be explained by the three variables: computational power (which is a proxy for mining difficulty), rate of coin production, and the relative hardness of the mining algorithm employed. This suggests that relative rates of production for given level of mining effort are paramount. For a given level of hashpower, increasing the difficulty will yield less units, and thus the relative cost of production. Similarly, reducing the block reward or employing a more rigorous mining algorithm will yield fewer units. In other words, this suggests that differences in the relative cost of production on the margin drive value formation for cryptocurrencies.

Using *Model B*, it is possible, in theory, to create an altcoin of high value simply by increasing its cost of production: choosing scrypt (or another even more difficult protocol) and reducing the coins produced per minute to some minuscule amount – this can be accomplished by increasing the Block Time and simultaneously reducing the Block Reward. Once that is achieved, the hard part is getting the computational power (and thus the mining difficulty) of the network up – and that is largely out of the control of the altcoin creator.

One important implication is that the total money supply, or ultimate number of units to ever be created is not a driving factor in value creation, rather it is the rate of unit creation that matters.

Of course, there are other subjective factors in determining the market price not included in the model, but which are yet to be identified. At any given point in time, any individual cryptocurrency may trade above or below its modelled value, the same as any other asset.

There is likely to be a speculative premium, as well as the tendency to hoard mined coins which will play an additional role in value formation, but which is more difficult to quantify and measure

2 The Decision to Mine for Altcoins and Miners' Arbitrage in Cryptocurrency Production

2.1 The decision to mine for altcoins

There exists efficient mobility of capital in switching mining effort from that of one coin to another; all one has to do is change the settings for the software or hardware to point the miner's hashing power towards mining another coin. Once those coins are mined and accumulated, they may be exchanged for bitcoin on any number of online exchanges.

Today, bitcoin is the stable equilibrium digital currency, and, for the most part, anybody wishing to transact in the real economy with a digital currency needs to use bitcoin. If obtaining bitcoins is the ultimate goal, a rational cryptocurrency miner would only direct mining effort at an altcoin if it provided for greater profitability than mining bitcoin directly over some period of time. What tends to happen is that any opportunities for excess profits are short-lived as competition drives all profit rates down to at least that of mining for bitcoin itself.

This apparent efficiency in removing opportunities to earn excess profits in mining seems to be the result of two forces: 1) competition of capital, as it is mobilized to mine for the more profitable coin it raises the aggregate network hashing power in that coin, causing the difficulty to subsequently increase. As the difficulty increases, profitability falls per unit of mining effort; and 2) the market exchange rate will change as mining participants actively produce and then sell relatively 'overpriced' coins.

Thus, both the bitcoin-denominated exchange price and the current difficulty of mining for the cryptocurrency in question relative to bitcoin's difficulty determines if there is an arbitrage opportunity, and acting on either variable will serve to eliminate that opportunity.

The baseline for profitability, then, or the regulating level of daily production, is the own-rate of return for bitcoin mining, measured in expected bitcoins per day per unit of mining power. For simplicity, I will peg that level of hashing power at a standard 1000 GigaHashes/sec (GH/s) of mining power, or 1 TeraHash/sec (TH/s). In practice, the actual hashing power of a miner is likely to deviate more or less from 1,000 GH/s, however this level tends to be a good standard of measure under current circumstances. The rate of bitcoin creation at the time of writing this paper is approximately 0.00831 BTC/day for every 1 TH/s of mining effort employed.⁷ The expected number of bitcoins expected to be produced per day can be calculated as follows:

$$\text{BTC/day}^* = [(\beta \cdot \rho) / (\delta \cdot 2^{32}) / \text{sec}_{\text{hr}}] \cdot \text{hr}_{\text{day}} \quad (1)$$

where: BTC/day* is the expected level of daily bitcoin production when mining bitcoin directly,

β is the block reward, ρ is the hashing power employed by a miner, and δ is the difficulty.^{8,9}

The constant sec_{hr} is the number of seconds in an hour, or 3600.

The constant hr_{day} is the number of hours in a day, or 24.

The constant 2^{32} relates to the normalized probability of a single hash per second solving a block, and is an attribute of the SHA-256 algorithm.

⁷ Given a current difficulty value for bitcoin of 60,813,224,039 and a block reward of 25 BTC.

⁸ Block reward is expressed in units of BTC/block

⁹ Difficulty is expressed in units of GH/block

The constants which normalize the dimensional space for daily time and for the mining algorithm can be summarized by the variable θ , which would equal $\theta = 24 \text{ hr}_{\text{day}} \cdot 232 / 3600 \text{ sec}_{\text{hr}} = 28,633,115.30667$. Equation (1) can thus be rewritten:

$$\text{BTC/day}^* = \theta (\beta \cdot \rho) / (\delta) \quad (2)$$

The only variables therefore are β , ρ and δ , and the hashing power, ρ , will be pegged to a fixed rate of 1,000 GH/s of hashing power.

An arbitrage opportunity exists when mining for any other cryptocurrency with the same amount of hashing power would produce a greater expected level of BTC/day than BTC/day*. To generalize equation (1) to account for any other altcoin, we simply introduce the current exchange rate of the altcoin/BTC pair, ε .¹⁰ Specifically, the market bid of the exchange rate is the price that matters since an arbitrageur would only be concerned with selling the altcoin to buy BTC.

Equation (3) indicates how many bitcoins would be obtained on average indirectly by mining for an altcoin instead:

$$\text{BTC/day}_{\text{altcoin}} = \theta (\beta_{\text{altcoin}} \cdot \rho) / (\delta_{\text{altcoin}}) \cdot \varepsilon \quad (3)$$

Under the no-arbitrage assumption that BTC/day* will be given as the own-rate of return for BTC, equation (3) can be re-arranged to solve for a theoretical equilibrium market price (of the bid) of the altcoin, holding the altcoin's difficulty constant:¹¹

$$\varepsilon^* = [\text{BTC/day}^*] / [\theta (\beta_{\text{altcoin}} \cdot \rho) / (\delta_{\text{altcoin}})] \quad (4)$$

If the altcoin's difficulty remains the same, there is a market opportunity for an arbitrageur to sell the relatively overpriced cryptocurrency until it reaches ε^* when exchanged for bitcoin on the market.

If, instead, the market price is held constant at ε^* , the difficulty can be thought of as relatively 'undervalued' and directing mining effort to that coin will produce excess profitability by subsequently exchanging those mined coins for bitcoin at price ε^* . Employing more mining power will necessarily increase the difficulty of that coin over time, so the arbitrage opportunity only exists until the difficulty is normalized and equilibrium is restored.

$$\delta^* = (\varepsilon \cdot \beta_{\text{altcoin}} \cdot \rho \cdot \text{sec}_{\text{hr}} \cdot \text{hr}_{\text{day}}) / (\text{BTC/day}^* \cdot 2^{32}) \quad (5)$$

Because equations (4) and (5) can be worked on by many different agents at the same time, arbitrage opportunities tend to be short-lived.

An example is useful here. A hypothetical individual miner has enough hashing power to earn 1 BTC/day*, on average. Alternatively, her same mining effort now could produce an expected 33,000 XYZ Coin per day, where XYZ Coin is a hypothetical altcoin that is traded against BTC on one or more exchanges. If the market bid is 0.00003996 BTC, she can exchange her XYZ and get in return: $33,000 \times 0.00003996 = 1.32 \text{ BTC/day}_{\text{altcoin}}$, making XYZ Coin mining right now 32% more profitable than mining bitcoin directly. As she and other miners continue to mine and subsequently sell their XYZ Coin, the market price in XYZ/BTC will fall as bids are cleared. The addition of new mining power in the XYZ network will also

¹⁰ All exchange rates are expressed in terms of Altcoin/BTC, or 1 Altcoin = X BTC

¹¹ BTC/day* also assumes Bitcoin's difficulty is constant during this period, not just that of the altcoin.

tend to make its difficulty rise, making it a more costly and less attractive alternative. It is worth noting that since there will tend to be orders of magnitude more mining effort directed at mining bitcoin than any altcoin at a given moment, while the new hashing power added to XYZ Coin may be a significant amount to XYZ Coin, the effort being removed from aggregate bitcoin mining is likely to be inconsequential and have no effect on bitcoin difficulty.

3 A Cost of Production Model for Bitcoin

3.1 Bitcoin production

As I have shown, the decision to mine for bitcoin comes down to profitability. A rational agent would not undertake production of bitcoins if they incurred a real ongoing loss in doing so. Bitcoin mining employs computational effort which requires the consumption of electricity to function, which must be paid for. This computational effort is directed at mining bitcoin, in competition with many other miners who presumably are also motivated by profit, on average. The more powerful the mining effort (the higher the hash-rate), the more likely it is to successfully mine bitcoins during a given interval (typically measured per day) for a given level of mining difficulty.

Therefore, success in finding bitcoins depends not only on the hashing power, but also on the difficulty level of the algorithm at the time that mining is undertaken. The difficulty specifies how hard it is to find a bitcoin during some interval, the higher the difficulty the more computational effort will be required to mine bitcoins at the same rate as with a lower difficulty setting. The bitcoin network automatically adjusts the difficulty variable so that one block of bitcoins is found, on average, every ten minutes. As more aggregate computational effort is added to mining bitcoins, the time between blocks will tend to decrease below ten minutes, the result being that the network will adjust the difficulty upwards to maintain the set ten minute interval accommodating the excess mining effort. Likewise, if mining effort is removed from the network, the length between blocks would grow longer than ten minutes and the network will adjust the difficulty downwards to restore the ten minute interval.

Each unit of mining effort has a fixed sunk cost involved in the purchase, transportation and installation of the mining hardware. It also has a variable, or ongoing cost which is the direct expense of electricity consumption. Each unit of hashing power consumes a specific amount of electricity based on its efficiency, which has a real-world cost for the miner. Because miners cannot generally pay for their electricity cost in bitcoin, they must refer to the currency price of a bitcoin to measure profitability given a real monetary cost of electricity.

It seems to be the case that the marginal cost of bitcoin production matters in value formation. Instead of approaching bitcoin as a digital money or currency, it is perhaps more appropriate to consider it a virtual commodity with a competitive market amongst producers.

The important variables in forming the decision to mine are: [1] the cost of electricity, measured in cents per kilowatt-hour; [2] the energy consumption per unit of mining effort, measured in watts per GH/s (or Joules per GH), a function of the cost of electricity and energy efficiency; [3] the monetary price of bitcoin in the market; and [4] the difficulty of the bitcoin algorithm.¹²

An individual would undertake mining if the marginal cost per day (electricity consumption) were less than or equal to the marginal product (the number of bitcoins found per day on average multiplied by the dollar price of bitcoin). If bitcoin production is a competitive com-

¹² The block reward also matters, but this value changes only after much longer intervals, approximately once every 4 years.

modity market, albeit a virtual one, then we would theoretically expect marginal cost to equal marginal product – which would also equal selling price.

The main cost in bitcoin mining is the energy consumption which is needed to facilitate the computational labour employed in mining.¹³ The actual market price is determined by the supply and demand for bitcoin at any given moment, while the cost of production might set a lower bound in value around which miners will decide to produce or not. While this lower bound could represent an intrinsic value, the actual observed price may deviate from that expected value for long periods of time, or may never converge to it.

Of course, there are likely to be many subjective motivations for bitcoin mining beyond the objective components elaborated in this paper. Individual decision makers may operate regardless of cost if they believe that there is enough speculative potential to the upside. Bitcoin mining may draw in those who find the features of anonymity and lack of governmental oversight attractive. Some miners may decide to hoard some or all of their lot and not regularly engage in offering mined bitcoins in the open market, a sort of bitcoin 'fetishism'. Some miners may be subject to an opportunity cost whereby it would be more profitable to expend the same electrical capacity for some other pursuit. Subjective rationales for mining may induce some individuals to make the decision to produce at a marginal loss for prolonged periods of time. The speculative and money-like properties of bitcoin, as a means of exchange and a potential store of value, add a subjective portion to any objective attempt at forming an intrinsic value. New and innovative uses of the bitcoin network for non-bitcoin specific applications are also likely to add value for mining.

3.2 The decision to mine for bitcoin and its cost of production

The objective decision to mine for bitcoins can be modelled. The necessary inputs are the dollar price of electricity, the energy consumption per unit of mining power, the dollar price of bitcoins, and the expected production of bitcoins per day which is based in part on the mining difficulty.¹⁴

Recall the model for determining the expected number of cryptocurrency coins to be mined per day on average given the difficulty and block reward (number of coins issued per successful mining attempt) per unit of hashing power, equation (1):

$$\text{BTC/day}^* = [(\beta \cdot \rho) / (\delta \cdot 2^{32}) / \text{sec}_{\text{hr}}] \cdot \text{hr}_{\text{day}}$$

The cost of mining per day, E_{day} can be expressed as:

$$E_{\text{day}} = (\text{\$price per kWh} \cdot 24 \text{ hr}_{\text{day}} \cdot \text{W per GH/s}) (\rho / 1,000) \quad (6)$$

Where E_{day} is the dollar cost per day for a producer, ρ is the hashpower employed by a producer, the $\text{\$price per kWh}$ is the price per kilowatt-hour, and W per GH/s is the energy consumption efficiency of the producer's hardware.

The marginal product of mining should theoretically equal its marginal cost in a competitive market, which should also equal its selling price. Because of this theoretical equivalence, and since cost per day is expressed in $\text{\$/day}$ and production in BTC/day , the $\text{\$/BTC}$ price level is simply the ratio of $(\text{cost/day}) / (\text{BTC/day})$. This objective price of production level, p^* ,

¹³ Other much smaller costs include internet service, hardware maintenance, computer cables etc.

¹⁴ For illustrative purposes only, the US dollar will be the currency used to price bitcoin. In reality, there are bitcoin miners worldwide, notably in Russia, Europe, and China who will buy electricity in their own regional currency and at their local rate.

serves as a logical lower bound for the market price, below which a miner would operate at a marginal loss and presumably remove them self from the network. p^* is expressed in dollars per bitcoin, given the difficulty and cost of production:

$$p^* = E_{\text{day}} / (\text{BTC}/\text{day}^*) \quad (7)$$

Note that p^* is a function of mining difficulty and the block reward in the denominator. Given an observed market price (p) and a known difficulty, one can solve for the break-even electricity cost per kilowatt-hour:

$$\text{\$price per kWh}^* = [p(\text{BTC}/\text{day}^*)/24\text{hr}_{\text{day}}] / W \text{ per GH/s} \quad (8)$$

Given a known cost of production and observed market price, one can solve for a break-even level of mining difficulty:

$$\delta^* = (\beta \cdot \rho \cdot \text{sec}_{\text{hr}} \cdot \text{hr}_{\text{day}}) / [(E_{\text{day}}/p) \cdot 2^{32}] \quad (9)$$

And, to solve for a break-even hardware energy efficiency, we can again rearrange terms given a market price, cost of electricity per kilowatt-hour, and difficulty:

$$W \text{ per GH/s}^* = (p \cdot \text{BTC}/\text{day}^*) / (\text{price per kWh} \cdot 24\text{hr}_{\text{day}}) \quad (10)$$

3.3 Discussion

These equations are useful in application as well as in theory. It informs miners objectively as to which price they should undertake or else give up mining. It also informs miners when to stop or start mining given changes in difficulty and electricity costs. Furthermore, looking at market prices for a given difficulty and known average electricity cost, the average energy efficiency of mining for the entire network can be imputed.

It is useful to consider a hypothetical example:

Assume that the average electricity cost for the world is approximately 13.5 cents per kilowatt-hour and the average energy efficiency of ASIC mining hardware currently deployed is 0.62 J/GH. The average cost per day for a 1 TH/s mining rig would be approximately: $(0.135 \cdot 24 \cdot 0.62) \cdot (1,000 / 1,000) = \$2.01/\text{day}$.

The number of bitcoins that same 1 TH/s of mining power can find in a day with a current difficulty of 60,813,224,039 is approximately 0.00831 BTC/day.

Because these two values (marginal cost and marginal product) are expected to be theoretically equivalent, to express them in dimensional space of \$/BTC we simply take the ratio: $(2.01 \text{ \$/day}) / (0.00831 \text{ BTC/day}) = \$241.877/\text{BTC}$.

This is surprisingly close to the current market value of around \$240-245 per BTC.

If the market price were to drop below that value, miners would be operating at a marginal loss and halt production. Continuing the analysis of this example, if the difficulty were to increase to greater than 61,404,400,615 holding all else constant, miners would cease operations. Also in this example, and holding all else constant, miners would cease operations if their energy costs rose to more than 13.6 cents per kilowatt-hour. Likewise, a miner would cease operations if their mining hardware consumed energy at an efficiency worse than 0.626 Watts per GH/s.

These figures are hypothetical for the purposes of elaborating the applicative usage of the equations introduced above, but have been chosen to be fairly close to current real-world practical, observed averages. It is worth noting the very small margins that exist for a variable to change and make mining for bitcoin no longer worthwhile for the average producer: for example electricity costs only need rise 0.01 cents or the difficulty by 1%.

As real-world mining hardware efficiency increases, which is a likely result of competition, the break-even price for bitcoin producers will tend to decrease. Low cost producers will compete in the marketplace by offering their product at lower and lower prices. Mining hardware energy efficiency has already increased greatly since the days of CPU or GPU mining. A research study found that the average mining efficiency over the period 2010-2013 was a staggering 500 Watts per GH/s (Garcia, et al., 2013). Today, the best ASIC mining rigs available for purchase have somewhere around 0.15 Watts per GH/s energy efficiency. The average energy efficiency right now across the mining network, which is the value which regulates the marginal cost, seems to be around 0.60-0.65 Watts per GH/s. This speaks to the rapid pace of technological advancement produced over the past few years and months in mining energy efficiency. The Bitcoin mining network is vast in size and scope and it is likely that some miners are at work with hardware that is older and less efficient than the best available.

Figure 1, below, illustrates how rapidly the energy efficiency of mining hardware for Bitcoin has improved over time. The rate of technological progress in this case has actually exceeded that predicted by Moore's Law.¹⁵

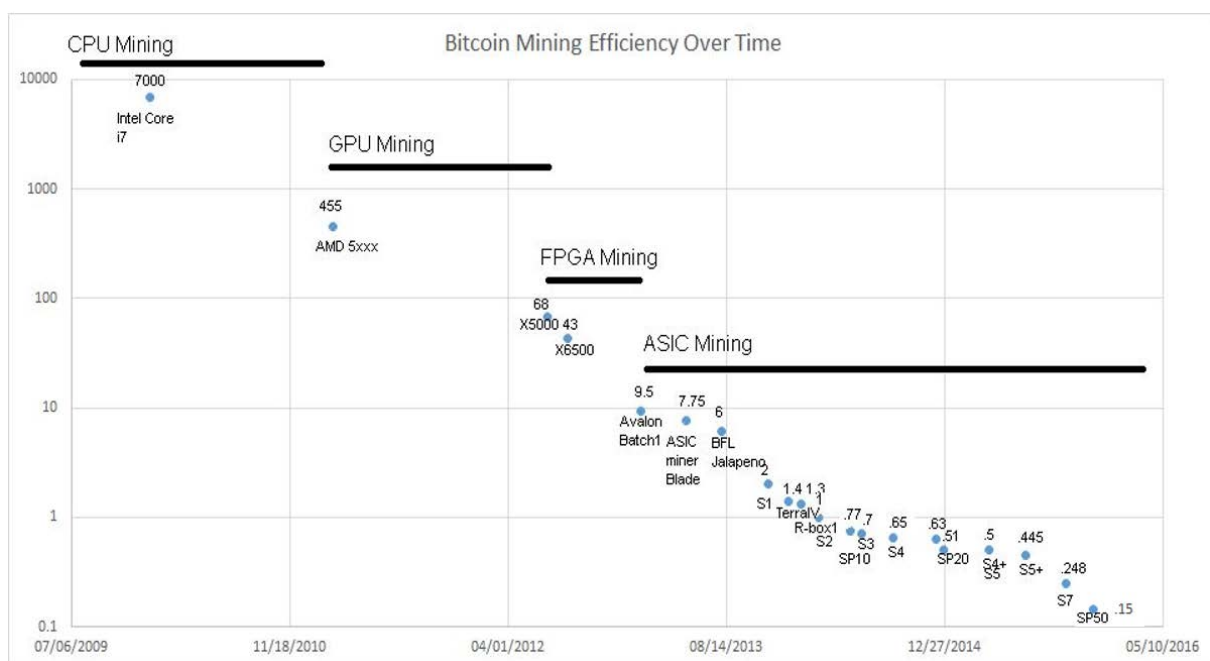


Figure 1: Bitcoin mining energy efficiency over time (log scale)

Bitcoin mining, unlike traditional commodity production, has the unique feature of a regular difficulty adjustment in order to maintain a steady rate of unit production over time – specifically, a block of bitcoins will be mined on average once every ten minutes, regardless of aggregate mining power. Unlike most produced commodities where the supply can change to accommodate fluctuations in demand, the supply of bitcoin is hardwired at its steady rate with the difficulty setting adjusting up and down to maintain that linear rate of production through time. In other words, the elasticity of supply is manifest in changes in the mining difficulty.

As energy efficiency increases, the difficulty adjustment acts as a stabilizing mechanism, increasing the cost of production; as more aggregate mining power is brought on line, the mining difficulty increases. For example, if a mining rig can find 1 BTC/day on average with to-

¹⁵ "Moore's law" is the observation that, over the history of computing hardware, the number of transistors in a dense integrated circuit, and therefore its processing power, has doubled approximately every two years.

day's difficulty, the same rig can expect to produce less per day when the difficulty increases 10% or 20% etc. If miners are not able to supply enough new coins to meet an influx of new demand, the market price can see increases while the cost of production remains largely the same. This would induce miners to increase their mining efforts which would then cause the difficulty to increase, raising the cost of production until presumably a new breakeven level is reached. This mechanism tends to counteract the downward tendency caused by increasing energy efficiency.

Figure 2, below, illustrates the relative change in mining efficiency compared to changes in mining difficulty over time. The left y-axis represents the inverse of the mining difficulty on a logarithmic scale, and is denoted by the dark blue line on the chart. The right y-axis is the mining efficiency, measured in joules per GH, and is denoted by the orange line. The x-axis is time from bitcoin's origin in 2009 to the present.

Initially, when bitcoin mining was only accomplished via a computer's central processor, or CPU, there were not many individuals involved with bitcoin mining, and the difficulty was very low. At the same time, mining was very inefficient. A computer's processor is designed to do many tasks such as run software and applications. It was discovered that a computer's graphics processor (GPU) was much better at solving the cryptographic algorithm used to mine for cryptocurrencies, and the difficulty grew rapidly as more mining power suddenly came on line. In Figure 2, the purple shaded areas indicate periods where the network size (difficulty) was increasing at a faster pace than technological change in mining efficiency. Green-shaded areas indicate periods where technological change has outpaced the growth of the network.

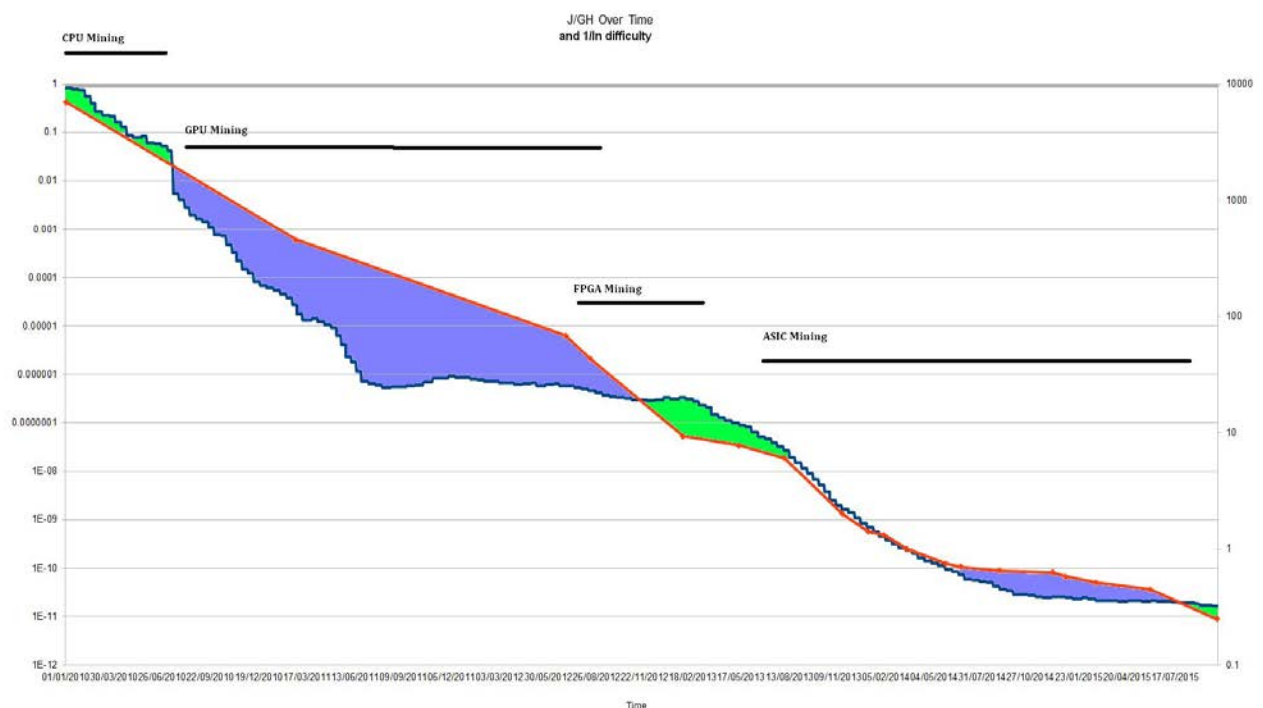


Figure 2. Bitcoin mining difficulty vs. mining energy efficiency over time (log scales)

GPU mining, while more efficient than CPU mining, was still not ideal. Video cards are designed for computer graphics and optimized for application such as gaming or design. As a result, mining with a GPU is not optimized for cryptocurrency mining. Furthermore, the manufacturers of graphics cards (and of CPUs for that matter) do not concern themselves with increasing the efficiency of their products to mine for cryptocurrencies. These manufacturers typically produce newer, better GPU hardware only when they can improve their primary

functionality. Therefore, there is no induced technological change to make these devices more efficient at mining even when the network size and mining difficulty is growing rapidly.

This all, however, changed with the introduction of application-specific integrated circuits, or ASICs, designed with the sole purpose to solve the encryption underlying cryptocurrencies. As a result, we begin to see marginal cost and marginal product begin to converge in mid- to late-2013 as they made their way to producers worldwide. Since then, there has been evidence of induced technological change, evidenced by the continued convergence of network size and mining efficiency since.

It is important to note that in the pre-ASIC period of bitcoin mining, the cost of production model outlined above would not hold. The capacity utilization of a CPU or GPU to mine for bitcoin is simply not efficient enough. One would not expect marginal cost to converge to marginal product when the hardware being used is not subject to competition. An apt comparison is that with ASICs it is like mining for gold with a pick and shovel – specifically made for such an activity – and when mining with CPUs/GPUs is like mining for gold with a shoe. While a shoe is not meant to mine for gold, one could conceivably collect some dirt in the shoe and find gold by happenstance. Just as the picks and shovels used for gold mining were induced to adapt and change, becoming steam shovels and later industrial mining operations, so too has bitcoin mining in the ASIC age seen such technological progress and consolidation due to competition.

One insight that could have sizable consequences for the cost of production of bitcoin relates to the block reward amount and how changes in this variable will impact BTC/day production. When bitcoin was launched, each block mined was composed of 50 bitcoins. That amount is set to halve every four years, and in 2012 the block reward became 25.¹⁶ The block reward will again halve to 12.5 bitcoins per block, expected mid-September, 2016, and will again in the year 2020, and so on. If we refer back to the illustrative example above and substitute a 12.5 BTC block reward for the current 25, the expected BTC*/day' becomes half of 0.00831, or 0.004155 per 1 TH/s. Using the hypothetical example above and given this new BTC*/day', the break-even price for a bitcoin would increase suddenly to \$483.75, holding all else constant.¹⁷ If the market price of bitcoin does not increase in turn, it will suggest that the breakeven efficiency has also increased at a more or less equivalent rate. This could have the effect of eliminating all but the most efficient producers all at once.

Conclusion

Beginning with a cross-sectional analysis to define the causes of relative value formation amongst cryptocurrencies, it was found that relative differences in costs of production on the margin are the main determinants. By looking at bitcoin-denominated relative prices, which are available on a number of online cryptocurrency exchanges, the high degree of price volatility found in the dollar-bitcoin exchange rate was eliminated. Cross-sectional analysis also was able to remove a number of other issues found in time-series analysis including any chance of non-stationary data or a small time horizon for the data set.

Next, using this result as a springboard, a series of equations were formalized to calculate how many units of a cryptocurrency a producer with a fixed amount of hashing power could expect to find, on average. Because Bitcoin is the stable equilibrium digital currency, even if some other altcoins are better or have various interesting features that Bitcoin lacks, it will be very difficult to dislodge. Therefore, the ultimate goal of any cryptocurrency producer operating in the real economy will to obtain bitcoins.

¹⁶ Lags in difficulty adjustment over time may result in the actual halving date occurring somewhat prior to or after 4 years.

¹⁷ The change in block reward will have no impact on difficulty. Rather, less BTC/day will be found given the same difficulty.

Given an efficient mobility of capital, a cryptocurrency producer will only mine for an altcoin if there is a greater profitability in that than using their equipment to mine for bitcoin directly. When these cases occur, markets tend to efficiently correct arbitrage opportunities ensuring that no altcoin is more profitable to produce than mining for bitcoin directly.

Finally, a cost of production model is put forward to establish break-even values for a bitcoin producer. Extrapolating that model to account of the average or regulating values for the aggregate Bitcoin mining network, the cost of production model can closely approximate the market price for Bitcoins versus dollars.

The implications are that cost of production drives value and anything that serves to reduce the cost of bitcoin production will tend to have a negative influence on its price. Increased mining hardware energy efficiency, lower worldwide electricity prices, or lower mining difficulty will all reduce the marginal cost of production. As mining efficiency increases due to technological progress, it lowers the cost of production and puts a negative pressure on the price. At the same time, the additional hashing power added to the global mining network will tend to increase the mining difficulty, and positively influence the price. The question will be which factor will outpace the other: technological progress (energy efficiency) or the size of the mining network (difficulty). A further implication is that when the Bitcoin block reward halves, it will effectively increase the cost of production overnight.

References

- Bergstra, Jan Aldert. *Bitcoin: not a currency-like informational commodity*. Informatics Institute, University of Amsterdam, (2014).
- Bouoiyour, Jamal, and Refk Selmi. *What Bitcoin Looks Like?* . No. 58091. University Library of Munich, Germany, (2014).
- Garcia, David, et al. *The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy*. Journal of The Royal Society Interface 11.99 (2014): 20140623.
- Gertchev, Nikolay. *The Moneyiness of Bitcoin*. www.mises.org (2013).
- Halaburda, Hanna, and Neil Gandal. *Competition in the Cryptocurrency Market*. Available at SSRN 2506463 (2014).
- Hanley, Brian P. *The false premises and promises of Bitcoin*. arXiv preprint arXiv:1312.2048 (2013).
- Harwick, Cameron. *Crypto-Currency and the Problem of Intermediation*. Available at SSRN 2523771 (2014).
- Nakamoto, Satoshi. *Bitcoin: A peer-to-peer electronic cash system*. Whitepaper. Consulted 8.2015 (2008): 28.
- Polasik, Michal, et al. *Price Fluctuations and the Use of Bitcoin: An Empirical Inquiry*. Available at SSRN 2516754 (2014).
- Van Alstyne, Marshall. *Why Bitcoin has value*. Communications of the ACM 57.5 (2014): 30-32.
- Woo, David, et al. *Bitcoin: a first assessment*. FX and Rates (2013).
- Yermack, David. *Is Bitcoin a Real Currency?*. No. w19747. National Bureau of Economic Research, 2013.
- Zhang, Yiteng, and Guangyan Song. *Economics of Competing Crypto Currencies: Monetary Policy, Miner Reward and Historical Evolution*. (2014).
- * Additional data sets collected from: coinmarketcap.com, coinwarz.com, cryptsy.com, bitcoinwisdom.com, and blockchain.info

Appendix

Dependent Variable: LOG(_PRICE)

Method: Least Squares

Date: 09/18/14 Time: 22:18

Sample: 1 83

Included observations: 66

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---------------------|-------------|-----------------------|-------------|--------|
| C | -9.526966 | 0.779882 | -12.21591 | 0.0000 |
| LOG(GH_S_HASHPOWER) | 0.685330 | 0.064390 | 10.64349 | 0.0000 |
| LOG(COINS_PM) | -0.983477 | 0.061908 | -15.88620 | 0.0000 |
| TYPE="Scrypt" | 7.461170 | 0.875470 | 8.522473 | 0.0000 |
| R-squared | 0.843083 | Mean dependent var | -5.493893 | |
| Adjusted R-squared | 0.835491 | S.D. dependent var | 3.888541 | |
| S.E. of regression | 1.577183 | Akaike info criterion | 3.807849 | |
| Sum squared resid | 154.2253 | Schwarz criterion | 3.940555 | |
| Log likelihood | -121.6590 | Hannan-Quinn criter. | 3.860287 | |
| F-statistic | 111.0381 | Durbin-Watson stat | 2.119851 | |
| Prob(F-statistic) | 0.000000 | | | |

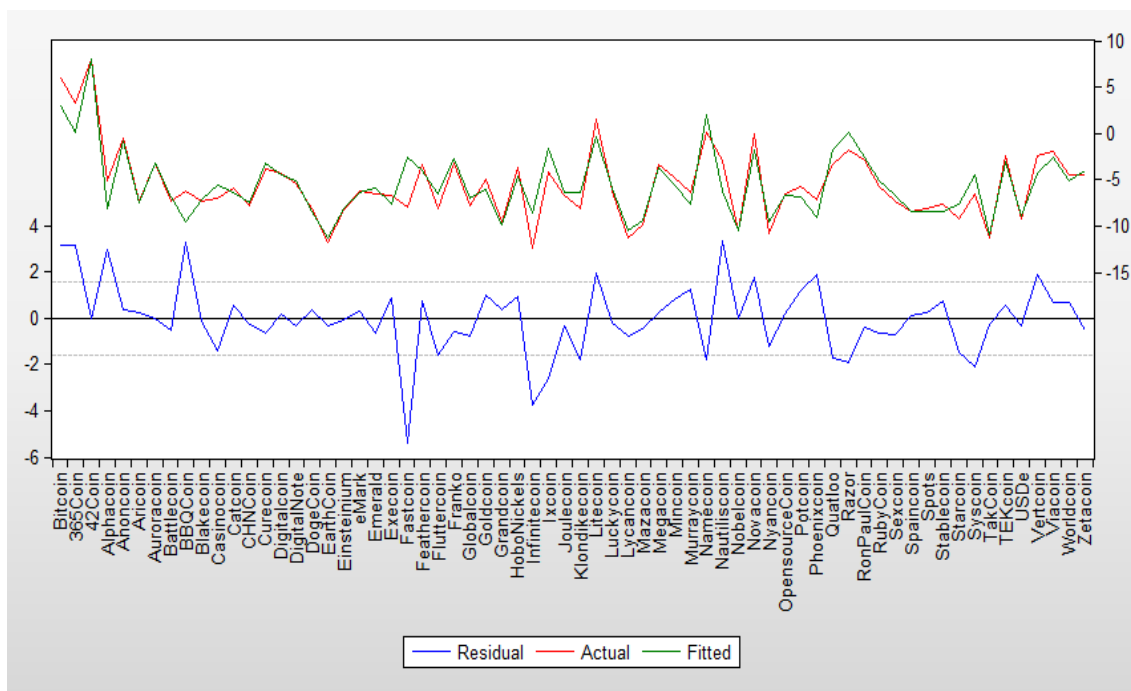
Pairwise Granger Causality Tests

Date: 12/11/14 Time: 10:11

Sample: 1 66

Lags: 2

| Null Hypothesis: | Obs | F-Statistic | Prob. |
|--|-----|-------------|--------|
| GH_S_HASHPOWER does not Granger Cause _PRICE | 64 | 65.5462 | 1.E-15 |
| _PRICE does not Granger Cause GH_S_HASHPOWER | | 0.01678 | 0.9834 |



HOW PERFECT OFFLINE WALLETS CAN STILL LEAK BITCOIN PRIVATE KEYS

Complete Research

Stephan Verbücheln, Humboldt-Universität, Berlin, Germany, verbuech@informatik.hu-berlin.de

Abstract

ECDSA has become a popular choice as lightweight alternative to RSA and classic DL based signature algorithms in recent years. As standardized, the signature produced by ECDSA for a pair of a message and a key is not deterministic. This work shows how this non-deterministic choice can be exploited by an attacker to leak private information through the signature without any side channels, an attack first discovered by Young and Yung for classic DL-based cryptosystems in 1997, and how this attack affects the application of ECDSA in the Bitcoin protocol.

Keywords: ECDSA, Bitcoin, Digital Signatures, Kleptography.

1 Introduction

Bitcoin is decentralized payment scheme first described in a publication by pseudonymous author Satoshi Nakamoto in (2008). Because its design is based on cryptographic protocols, the term *cryptocurrency* has become common to refer to systems like Bitcoin and its variants. In a nutshell, users hold public-private key pairs, where the public keys function as an account numbers, and the private keys enable them to spend money from those accounts.

Bitcoin has become very popular and commercially adopted in recent years. The exchange rate has risen to a point that the value of all coins in circulation adds up to \$5 billion as of November 2014 (peaking even higher before). Because of the value of Bitcoin assets, it has become more and more important for users – especially large-scale users like online shops – to protect their private keys, as stealing a Bitcoin private key enables an adversary to steal their money.

Many ideas have come up on how to keep the user's private keys secret. The regular Bitcoin wallet software for PC operating systems encrypts the private keys using AES and only decrypts them when the user wants to create a transaction. This gives basic protection against PC malware for the time the user does not use his private key. But when the user wants to generate a transaction and enters his passphrase, the unencrypted keys are in the application memory, and thus are not protected against malware.

The next step is to use a dedicated PC for Bitcoin only, and keep it disconnected from the Internet. This prevents the PC from being infected with malware, and even in case it gets infected, the malware has a hard job to send the secret keys home to the attacker. As an even further step, some companies start producing special purpose hardware which is designed to manage keys and create transactions, but to never release the keys themselves.

The strength of the latter two settings comes from the fact that even if the device with the Bitcoin private keys on it is malicious, it cannot send any private information to attackers as long as the user makes sure that only transactions leave the device. This is supposedly achieved, if the device for example writes the transaction into a file, and the user copies the file using a portable storage medium.

But, as we will see later, even if the user actually ensures that only the transaction leaves the wallet device, there is still a problem: ECDSA. Bitcoin transactions contain ECDSA signatures. In the process of creating

a ECDSA signature, like in the classic DSA, the creator has to choose a random number. Adam Young and Moti Yung showed in (1997), how this random number can be used by a malicious implementation of DSA to leak private information. This can be done in such a way that only the attacker, who made the malicious implementation, can extract the secret from the signature, and that the distribution of malicious signatures is polynomially indistinguishable from legitimate signatures by anybody else.

This paper will outline how this attack can be used for ECDSA as well, and what security issues arise from that for applications of ECDSA like Bitcoin.

2 Related Attacks

A number of attacks on Bitcoin keys have been around in the wild as well as in research publications, e.g. in works by Nicolas Courtois and Filippo Valsorda (Valsorda 2014), (Courtois, Emirdag and Valsorda 2014). Those attacks are enabled by same properties in the ECDSA scheme as the attack presented in this paper: When generating a ECDSA signature, the user has to generate a new random number that has to be kept secret, because knowledge of this number enables an attacker to compute the private key.

The attacks outlined by Valsorda and Courtois (and successfully run by many attackers of unknown attackers on the actual Bitcoin network) make use of weak randomness. If a user's wallet software is reusing the same random number for generating signatures several times, attackers can compute the private key and steal their Bitcoins.

The attack outlined in this paper is different, because it does not make use of bad randomness, which would potentially be detected by attackers scanning the network. In this attack, a malicious wallet implementation is generation two of those random numbers in a way that the attacker can compute the private key from two signatures, while for everybody else (who does not have the attacker's private key), the signatures look random.

3 Definitions

In our considerations, the *user* is the person, who wants to use a publicly specified cryptosystem, e.g. ECDSA. The *attacker* is the person, who creates a malicious implementation which is used by the user. The intuition is that the attacker's implementation can differ from how the cryptosystem is specified, but that the inputs and outputs have to comply with the specification. In addition we don't want the user or any third party to be able to distinguish the outputs generated by the malicious implementation from the outputs generated by a specification-compliant implementation.

3.1 Kleptographic SETUP

We will use the definition of kleptographic setups given by Young and Yung in (1997).

Definition. Let C be the black-box implementation of a cryptosystem with publicly known specification. A kleptographic (*regular*) *SETUP* (Secretly Embedded Trapdoor with Embedded Protection) is a modified algorithm C' such that:

1. The input of C' agrees with the public specifications of the input of C .
2. C' computes efficiently using the attacker's public encryption function E (and possibly other functions as well), contained within C' .
3. The attacker's private decryption function D is not contained within C' and is known only by the attacker.
4. The output of C' agrees with the public specifications of the output of C . At the same time, it contains published bits (of the user's secret key) which are easily derivable by the attacker (the output can be generated during key generation or during system operation like message sending).

5. Furthermore, the output of C and C' are polynomially indistinguishable (as in Goldwasser, Micali 1984) to everyone except the attacker.
6. After the discovery of the specifics of the setup algorithm and after discovering its presence in the implementation (e.g. reverse-engineering of hardware tamper-proof device), users (except the attacker) cannot determine past (or future) keys.

4 A Kleptogram for Elliptic Curves

In this section, we will see how elliptic curves can be used to hide information in the choice of a random number like it is performed in ECDSA. This will be the central building block in our later attack.

Given an elliptic curve E . Let G be a point on E of order n . Let d be the attacker's private key and $Q = dG$ the corresponding public key. Let \mathcal{R} be a cryptographically strong pseudo-random number generator with hidden seed. Without loss of generality, we assume that \mathcal{R} outputs a value less than n . Let α, β, ω be fixed integer constants, with ω being odd.

4.1 Generation

The malicious implementation of a elliptic-curve cryptosystem generates two subsequent choices c_1, c_2 the following way:

First round.

1. Pick random $c_1 < n$.
2. Compute $M_1 = c_1G$.
3. Store c_1 in non-volatile memory.
4. Output M_1 .

Second round.

1. Pick random bit $t \in \{0, 1\}$.
2. Compute $Z = (c_1 - \omega t)G + (-\alpha c_1 - \beta)Q$.
3. Compute $c_2 = \mathcal{R}(Z)$.
4. Compute $M_2 = c_2G$.
5. Output M_2 .

4.2 Recovery

Now the attacker is able to extract the (secret) value of c_2 from M_1, M_2 and his private key d as follows:

1. Compute $R = \alpha M_1 + \beta G$.
2. Compute $Z_1 = M_1 - dR$.
3. If $M_2 = \mathcal{R}(Z_1)G$ then output $c_2 = \mathcal{R}(Z_1)$.
4. Compute $Z_2 = Z_1 - \omega G$.
5. If $M_2 = \mathcal{R}(Z_2)G$ then output $c_2 = \mathcal{R}(Z_2)$.

4.3 Indistinguishability

Values in the first round are chosen at random. An attacker wants the distribution of values from the second round to be indistinguishable from values generated at random like in the first round.

For this reason, the second round makes use of a seeded pseudo-random number generator. To make the distribution of the generator's output as hard to distinguish as possible, the attacker wants the number of potential values for Z as large as possible, ideally n different values.

The attacker can achieve this by tweaking the constants α , β and ω such that

$$G_1 = (-d\beta - \omega)G$$

$$G_2 = (-d\beta)G$$

$$G_3 = (1 - d\alpha)G$$

have preferably high orders, the ideal being n . In that case, Z and thus $c_2 = \mathcal{R}(Z)$ can take on any value up to n . Assuming that the output of \mathcal{R} is polynomially indistinguishable from random numbers of equal distribution, we can follow that only the attacker himself can distinguish the two (by blindly applying his recovery procedure).

5 ECDSA and Bitcoin

In this section, we will see how the above construction can be used to create a SETUP for ECDSA.

5.1 Setup for ECDSA

Let's recall ECDSA. For an elliptic curve E with a generator point G of order n , an ECDSA user's private key is a number $d < n$. The public key is the corresponding curve point $Q = dG$. Let \mathcal{H} be the cryptographic hash function used.

A signature for a message m is generated as follows:

1. Pick a random value $k < n$.
2. Compute $R = (r_1, r_2) = kG$.
3. Compute $r = r_1 \bmod n$.
4. Compute $s = k^{-1}(\mathcal{H}(m) + dr) \bmod n$.
5. Output $\sigma = (r, s)$

The signature is verified by another party as follows:

1. Compute $R' = (r'_1, r'_2) = s^{-1}\mathcal{H}(m)G + s^{-1}rQ$.
2. Compute $r' = r'_1 \bmod n$.
3. Accept the signature if and only if $r = r'$.

Note that a number k_i is chosen by the algorithm for each signature σ_i . A black-box implementation of ECDSA can substitute the values k_i, k_{i+1} in two consecutive signatures σ_i, σ_{i+1} by c_1, c_2 generated like in section 3. This enables the attacker to extract the value of $k_{i+1} = c_2$ from $\sigma_1 = (r_1, s_1), \sigma_2 = (r_2, s_2)$ as described, because r_1, r_2 are the equivalent to m_1, m_2 respectively from section 3.

Now note that knowledge of the value k for a single signature σ already enables an attacker to compute the private key d , since from step 4 in the signature generation, it follows:

$$d = (\mathcal{H}(m) - sk)r^{-1} \bmod n$$

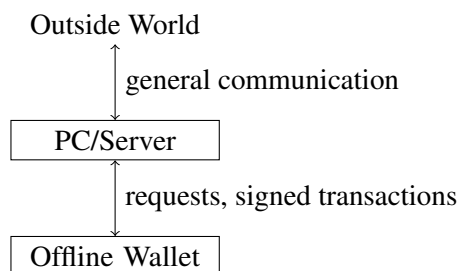


Figure 1. Bitcoin configuration with an offline wallet

5.2 Limitations

One limitation with this attack is that the attacker has to know in advance which curve and which generator point the user will choose to set up his implementation. Otherwise he cannot generate his own public-private key pair. This was a big limitation, when Young and Yung (1997) published the attack for classic Diffie-Hellman settings, because it was common that each user generates his own new group \mathbb{Z}_p and a generator $g \in \mathbb{Z}_p^*$ for each application.

With elliptic curve cryptography, this limitation became less relevant. Since it is more expensive to generate a new strong pair of a curve and a generator point from scratch, elliptic curves are generated and analyzed by scientists, and only a few elliptic curves are published (with corresponding parameters including the generator point) by standardization bodies such as the U.S. governments' NIST. Some ECDSA applications even specify a particular single curve to be used, e.g. Bitcoin uses only the secp256k1¹ curve for ECDSA.

5.3 Bitcoin

Bitcoin is a decentralized payment system, which is implemented as a peer-to-peer network (Nakamoto 2008). In order to pay money to one another, users create little files named *transactions*. These transactions are broadcasted into the network and collected in a large log file called *blockchain*, which is maintained by all participating peers in a cooperative manner.

For our considerations we don't have to go further into how the blockchain is maintained in detail. We just have to keep in mind that the blockchain collects all valid transactions, and that the previous blockchain history is what gives the user a balance that he can spend.

To participate in the payment and receive money, a user creates a Bitcoin *address*. He creates an ECDSA public-private key pair.² To generate the address, the public key is hashed twice using the SHA256 algorithm.

Now a standard transaction is a file that specifies a number of inputs and outputs. For inputs, the transaction refers to a previous transaction that has charged one or more of the user's addresses with coins. For outputs, the transaction specifies one or more Bitcoin addresses owned by the receiver. To make the transaction valid, it has to satisfy the following conditions:

- The specified inputs must have a higher or equal value than the outputs are receiving.
- The transactions specified in the inputs must be part of the blockchain and not yet spent.
- The user has to prove that he owns the addresses from the input transactions by signing the transaction with the corresponding private keys.

¹ Published in 2000 by Certicom Research.

² Note that the secp256k1 curve is the only elliptic curve used in Bitcoin.

The last condition is the one that makes sure that only the owner of an address can spend the received coins, because only he knows the ECDSA private key.

Now note that since the signatures in a transaction are plain ECDSA signatures, the kleptographic SETUP from section 4 can just be applied in a straight forward manner. This means that a malicious programmer or hardware designer can implement a Bitcoin wallet in a way that leaks the secret key without any side channels using only two signatures. Note that since there might be more than one signature in a transaction, this can happen in a single transaction if two inputs associated with the same address are used. Nobody but the attacker is able to distinguish such a malicious transaction from a normal one.

5.4 Deterministic Choice of k

The kleptographic attack on ECDSA is very easy, because the value k has to be chosen during the signature creation. As we have seen, k has to be secret to prevent adversaries from extracting the secret key d . Only $R = kG$ is published with the signature.

Another slightly related security issue also arose from the fact that k has to be chosen by the signature algorithm. If two values k_1, k_2 in two different signatures have a known linear relationship $k_2 = ak_1 + b$ with $a, b \in \mathbb{Z}$, the private key d can be extracted from the two signatures without the knowledge of the values k_1, k_2 , since it results in two linear equations with only d and k_1 unknown.

Because of this known attack, some proposals have been made on how to choose the number k deterministically, e.g. RFC6979 or the specification of EdDSA by Bernstein, Duif, Lange, Schwabe, Yang (2012). These proposals generate k deterministically using a cryptographic hash function with the message and the user's private key as inputs. The result is that the same user always creates the same signature for a given message. Note that it is crucial here to have the private-key part of the input. If k would be derived from the message alone, it would be public information and therefore useless.

This measure gives only limited protection against the kleptographic attack since it can only be verified using the private key. The whole point of a dedicated Bitcoin wallet is that the user wants to make sure that the private keys never get anywhere outside of it, which means that even the user himself cannot verify the signature using a second computer.

6 Potential Solutions

One potential counter-measure could be the following:

1. The signature device generates the signature with deterministic k as specified in RFC6979.
2. In addition to that, the device delivers a zero-knowledge proof that k was indeed generated as specified.

This is possible since it is known that zero-knowledge proofs exist for any NP statement. But with this solution, a new problem arises. The whole point of using an offline Bitcoin wallet is that it does not leak any information into the public except the legitimately generated transactions. If we let the device output a zero-knowledge proof in addition to that, this proof may introduce new ways for the adversary to leak information. This means that the zero-knowledge protocol for this application has to be chosen carefully. Another counter-measure would be to strictly not use any address more often than once. Although this way of using Bitcoin addresses is recommended already because of privacy considerations, there exist some use cases where this may not be feasible. For example, a public donation address for a charitable foundation is supposed to be used by multiple donors over and over again for a long period of time. In order to transfer the donations, a signature has to be created for each incoming payment.

A deterministic choice of k alone (as described in the last section) would not help much either, because knowledge of the private key is necessary to verify that k has indeed been chosen as specified. But there are still two advantages arising from using a deterministic method: First, it limits the choice of k if the user signs the exact same document twice, as k has to be identical given the same message and private key.

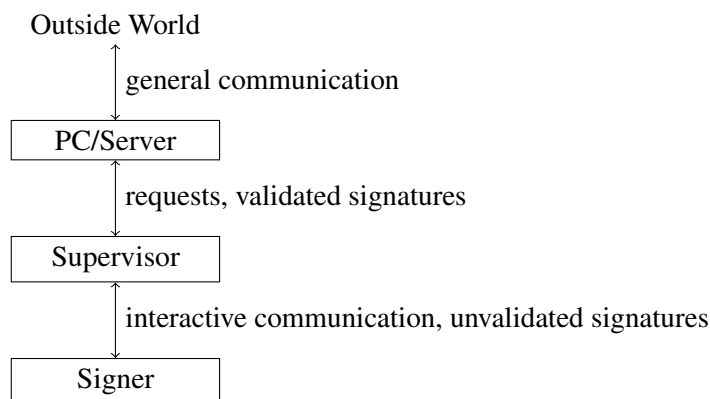


Figure 2. Configuration with a supervisor

Secondly, it would enable the user to detect malicious signatures later, for example after a key has expired and it is safe to transfer the private key to another computer.

6.1 Interactively Generate k

The ability of the signer to leak information through his choice of k come from the fact that he is allowed to choose k freely. If an outside agent could force the signer to choose k from the equal distribution, that wouldn't be possible.

A common way to solve this cryptographic problem is an interactive protocol, where two parties choose a common random string. Such a protocol is arranged in a way that neither of the two parties can influence the resulting random string to his wishes, because both parties have to make their choice for their part before the other one's choice is revealed. This can be enforced using a commitment scheme.

In a blogpost³ on firmcoin.com, the Certimix company describes such an interactive protocol that deals with the problem. In addition to the signing device (i.e. the offline wallet), we have a supervisor device that checks whether the signing is done properly. The protocol works as follows:

For an elliptic curve E with a generator point G of order n be $d < n$ the private key and the corresponding curve point $Q = dG$ the public key. Let \mathcal{H} be the cryptographic hash function used.

A signature for a message m is generated as follows:

1. Supervisor and signer generate k together:
 - a) The signer picks a random value $t < n$.
 - b) The signer computes $T = tG$.
 - c) The signer computes $h_T = \mathcal{H}(T)$.⁴
 - d) The signer sends h_T to the supervisor.
 - e) The supervisor picks a random value $u < n$.
 - f) The supervisor sends u to the signer.
 - g) The signer sends T to the user.
 - h) The supervisor verifies that $h_T = \mathcal{H}(T)$.
 - i) The signer computes $k = t \cdot u \pmod n$.
2. The signer computes $R = (r_1, r_2) = kG$.
3. The signer computes $r = r_1 \pmod n$.

³ Blogpost from June 20, 2013, "No subliminal Channel", <http://firmcoin.com/?p=52>

⁴ This is a commitment to T .

4. The signer computes $s = k^{-1}(\mathcal{H}(m) + dr) \pmod n$.
5. The signer sends the resulting signature $\sigma = (r, s)$ to the supervisor.
6. The supervisor releases the signature after verifying that k was generated correctly:
 - a) The supervisor computes $R' = (r'_1, r'_2) = uT$.
 - b) The supervisor verifies that $r = r'_1 \pmod n$ holds.

Note that only the supervisor has to communicate with the outside world. But there are still some things that we have to be aware of: Even though neither the signer nor the supervisor can manipulate the choice of k/r as published with the signature, the signer might still leak information to the supervisor through his choices of t_i . There are still advantages coming from the protocol: As the supervisor does not have to store any critical data in persistent memory, the security properties are different. In addition, as long as either one of signer and supervisor works correctly, the signature cannot leak information through the choice of k .

This protocol has the disadvantage that it requires a lot of interaction. But we can make this protocol more practical by performing a prearrangement step where a list of random numbers is generated in advance. Let E again be an elliptic curve with a generator point G of order n , be $d < n$ the private key and the corresponding curve point $Q = dG$ the public key. Let \mathcal{H} be the cryptographic hash function used.

Prearrangement

1. The signer chooses a list of random numbers t_1, \dots, t_ℓ with $t_i < n$ for each $i \leq \ell$.
2. For $i \leq \ell$, the signer computes $T_i = t_i G$ and $h_{T_i} = \mathcal{H}(T_i)$.
3. The signer stores the list of choices t_1, \dots, t_ℓ for later use.
4. The signer sends the list of hashes $h_{T_1}, \dots, h_{T_\ell}$ to the supervisor.
5. The supervisor stores the list of hashes.

Signing

1. The supervisor generates a random number $u < n$.
 2. The supervisor sends (m, u) to the signer.
 3. The signer computes $k = ut_1 \pmod n$ and $T_1 = t_1 G$, where t_1 is the first element from the list of choices.
 4. The signer generates the ECDSA signature $\sigma = (r, s)$ using k as the nonce.
 5. The signer sends (σ, T_1) to the supervisor and removes t_1 from the list of choices.
 6. The supervisor computes $R' = (r'_1, r'_2) = uT_1$.
 7. The supervisor verifies that $r = r'_1 \pmod n$ and $h_{T_1} = \mathcal{H}(R')$.
 8. If successfully verified, the supervisor publishes the signature and removes h_{T_1} from the list of hashes.
- If any verification step fails, the supervisor should cancel the protocol and alert the user.

7 Conclusion

Without a satisfying solution, there is only one conclusion to draw from this problem: Users cannot trust any implementation of ECDSA or Bitcoin, which they cannot fully verify.

Note that this does not only affect strict black-box implementations such as closed-source programs. A user with high security requirements (like an online shop that wants to accept Bitcoin payments on a large scale) would have to use an implementation, which can be verified by his own staff (or at least a reliable partner company). A hard-to-read implementation like OpenSSL may be insufficient, because the particular implementation of ECDSA may be hard to verify. Such implementations feature many

variations on the pure algorithm to improve performance (e.g. using CPU-dependent assembly language) or to harden the implementation against timing attacks). In addition, it is hard to verify in a large program like OpenSSL, which code is actually executed when you perform a certain operation.

A similar problem arises with embedded cryptographic chips like smartcards. Such devices are designed to never release the private keys and to make it hard for an outside analysis to read out secret data. The fact that the leak in the kleptographic attack is so well hidden makes it hard for a chip manufacturer to prove to the customers that the device does not leak any secret information in ECDSA signatures.

The paranoid among users would even have to compile the program themselves to be sure that the code they are reading really matches the code they are running. To verify that the executed code actually matches the source code is even harder for small embedded devices (like dedicated Bitcoin wallets or crypto smartcards) than in the setting of an offline PC.

References

- Daniel J. Bernstein, Niels Duif, Tanja Lange, Peter Schwabe, Bo-Yin Yang, “High-speed high-security signatures”, *Journal of Cryptographic Engineering*, September 2012, Volume 2, Issue 2, pp 77-89, 2012, Springer-Verlag
- Shafi Goldwasser, Silvio Micali, “Probabilistic Encryption”, *Journal of Computer and System Sciences*, Volume 28, Issue 2, April 1984, Pages 270–299, 1984, Elsevier
- Satoshi Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System” (pseudonymously published in 2008)
- Internet Engineering Task Force (IETF), “Request For Comments (RFC) 6979”
- Adam Young, Moti Yung, “The Dark Side of “Black-Box” Cryptography or: Should We Trust Capstone?”, *Advances in Cryptology — CRYPTO '96*, Lecture Notes in Computer Science Volume 1109, 1996, pp 89-103, 1996, Springer Berlin Heidelberg
- Adam Young, Moti Yung, “The Prevalence of Kleptographic Attacks on Discrete-Log based Cryptosystems”, *Advances in Cryptology — CRYPTO '97*, Lecture Notes in Computer Science Volume 1294, 1997, pp 264-276, 1997, Springer Berlin Heidelberg
- Filippo Valsorda, “Exploiting ECDSA Failures in the Bitcoin Blockchain”, *talk at HITB 2014*
- Nicolas T. Courtois, Pinar Emirdag, Filippo Valsorda, “Private Key Recovery Combination Attacks: On Extreme Fragility of Popular Bitcoin Key Management, Wallet and Cold Storage Solutions in Presence of Poor RNG Events”, *Cryptology ePrint Archive: Report 2014/848*, IACR

Trust in Digital Currency Enabled Transactions Model

Completed Research

Alex, Zarifis, University of Mannheim, Mannheim, Germany, azarifis@mail.uni-mannheim.de

Xusen, Cheng, University of International Business and Economics, School of Information Technology and Management, Department of E-Business, Beijing, China, xusen.cheng@gmail.com

Salomi, Dimitriou, University of Nicosia, UNIC Online, Nicosia, Cyprus, s.dimitriou@unic-online.com

Leonidas, Efthymiou University of Nicosia, UNIC Online, Nicosia, Cyprus, l.efthymiou@unic-online.com

Abstract

This research extends theories of trust from e-commerce to incorporate digital currencies. In particular trust in business to consumer e-commerce transactions carried out using digital currencies such as Bitcoin is explored. A model of online trust is considered to be valid in this different transaction context but the significance of each construct changes and some extensions are necessary. The role of institutional trust in transactions has differences that are explored and new constructs are suggested. These new constructs are incorporated into a new digital currency enabled transactions trust model. The results support the validity of the role of the rate of adoption and reputation of digital currencies as part of situational normality. The nature of the digital currency itself, the digital currency payment system, the payment intermediary, the digital currency P2P infrastructure, self-imposed and external regulation are also considered valid as part of structural assurance. These findings can be used by those developing the related technology, the vendors and regulatory institutions to increase consumer trust in digital currency enabled transactions in order to extend adoption and use.

Keywords: Digital currency, trust, Bitcoin, e-commerce, business to consumer.

1 Introduction

While there are a number of perspectives on what constitutes a digital currency the most widely adopted view which is also the one used in this research is that they allow the transfer of value online without a traditional bank being involved (Hett, 2008). The terms cryptocurrencies and virtual currencies are also used but they have a technology and virtual world bias respectively. Current research into digital currencies can be separated into three threads. The main focus of the media, practitioners and research (Sapuric and Kokkinaki, 2014; Hayes 2015) has been on the value of one such digital currency Bitcoin, considering this an important indicator of these currencies success. A second research thread is developing the underlying technology (Gjermundrod and Dionysiou, 2014; Verbücheln, 2015). A third thread, which this research is part of, attempts to explore and better understand the inherent functionality and value of this technology (Giaglis and Kypriotaki, 2014; Zarifis et al. 2014) considering the functionality and value as the deciding factors in their adoption in some shape or form (Internal Revenue Service, 2014). While the current price of Bitcoin depends on the actions of certain organizations and individuals the inherent characteristics are not and may therefore be a better medium and long term predictor of adoption.

This research attempts to better understand consumer trust in transactions using digital currencies so that our understanding of one of the inherent characteristics of this technology is developed. The consumer's perspective and level of trust in a technology is an important factor in the level of its adoption particularly when there is some financial risk. This paper develops previous research (Zarifis et al. 2014) which was more exploratory using qualitative methods by applying a quantitative data collection and analysis. The quantitative data collection was guided by the previous findings (Zarifis et al. 2014).

2 Digital Currency Enabled Transactions

There are a number of digital currencies in addition to Bitcoin including Litecoin, Ripple, Anything Point, Facebook Credits, Amazon Coin and Linden dollars. This research chose Bitcoin to explore but the findings should be relevant to other digital currencies also. Bitcoin, like Litecoin and Ripple, is both a currency and a transaction system (Nakamoto, 2009). Furthermore these three currencies are not limited to a specific environment (Akins et al. 2013) in the way Linden dollars can only be used in Second Life. For the consumer to use Bitcoin they need to have an internet connection, a user version of the Bitcoin software known as Bitcoin wallet and another user such as a retailer willing to carry out a transaction. The transaction is either at a minimal fee or no cost. The significantly lower cost in comparison to transactions implemented by traditional banks is one of the main factors that is attractive to consumers.

When the consumer considers collaborating online trust is an important factor. Trust becomes an even more decisive factor when making a transaction online (Jarvenpaa et al. 1998; O'Brien, 2000; Bhattacharjee 2002). Trust is considered to have two components trusting beliefs and trusting behavior (McKnight et al. 1998; Pearce, 2007). In addition to this psychological dimension of trust there is also a more sociological dimension that includes institution based trust. Institution based trust refers to the trust in the institutions involved in the transaction such as a regulator, a government or a bank card provider (McKnight et al. 1998).

Trust in digital currency enabled transactions includes the issues related to trust in online transactions. Therefore the research in the area of online trust should be applicable. In addition to being an online transaction however it has some different characteristics and a partly different context due to the use of

digital currencies and therefore warrants particular attention. Previous research (Zarifis et al. 2014) attempted to explore trust in this area and adapt an existing model to this different context.

3 Digital Currency Trust Model

The model proposed and explored in the first qualitative stage of this research (Zarifis et al. 2014) is a development of a widely adopted trust model (McKnight et al. 2002) and shows constructs of trust and their relationships. The model combines constructs applicable to trust in all online transactions such as the general web experience (McKnight et al. 2002) with constructs specific to digital currencies such as the digital currency payment system (Zarifis et al. 2014).

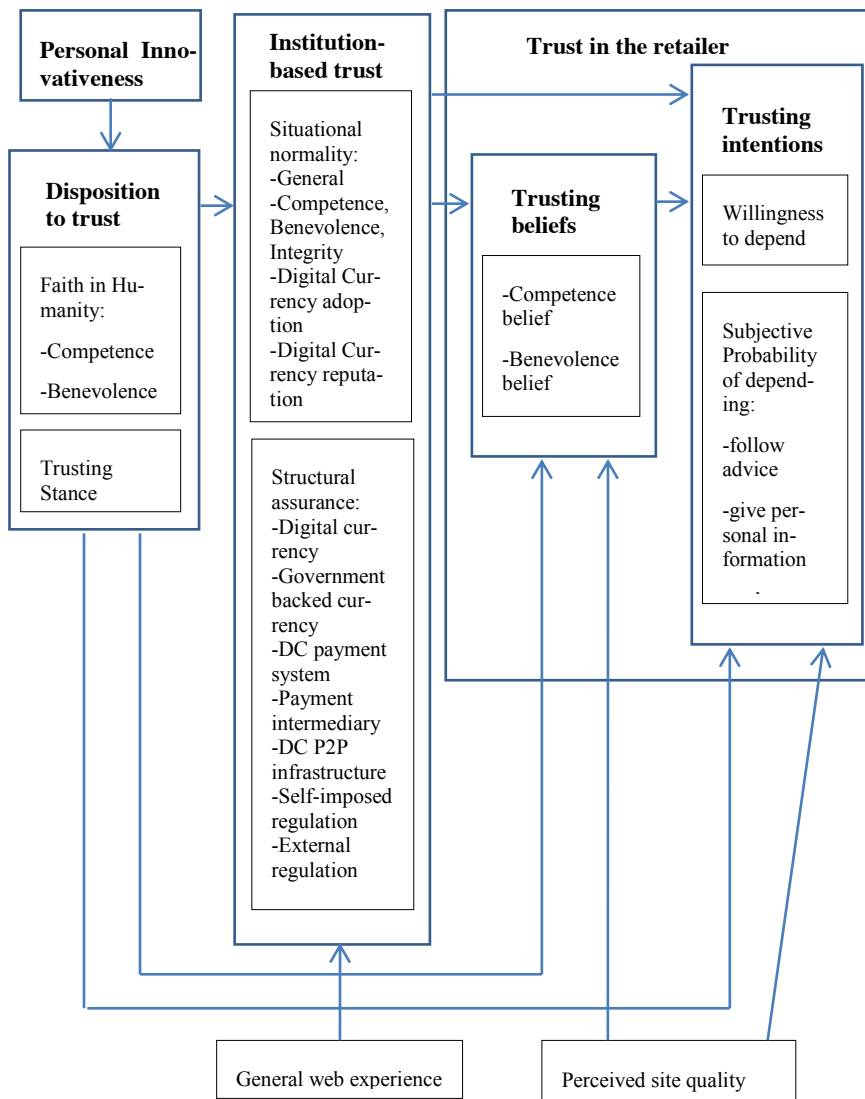


Figure 1. Digital currency enabled transactions trust model (Zarifis et al. 2014), an extension of the web trust model (McKnight et al. 2002).

The overarching constructs of the model personal innovativeness, disposition to trust, institution based trust, trusting beliefs, trusting intentions, general web experience and perceived site quality remain the same. These overarching constructs along with the original model have been validated extensively and are therefore a strong basis to develop and extend into the different context. These constructs had been first used in other disciplines before being applied to offline business to consumer commerce. They were then extended to the online context in a similar way to which they are being extended here to a partially different context.

Previous exploratory research (Zarifis et al. 2014) suggested that the constructs of personal innovativeness, disposition to trust, trusting beliefs, trusting intentions, general web experience and perceived site quality remained unchanged and valid for digital currency enabled transactions. These constructs are related to the psychology of the consumer and the online context so it is reasonable for them not to change as they are similar in the different context of digital currencies.

The differences identified to be explored further were in the area of institution based trust both in situational normality and structural assurance. The sociological dimension of institutional trust is posited to be significantly different to online transactions without digital currencies. Institutional trust is defined as the environment and institutions that shape and influence it (McKnight et al. 1998). For digital currency enabled transactions the environment is the internet. While the role of institutional trust has been researched extensively it is of particular interest for digital currencies because institutions that shape them are very different to the government organizations and private companies that shape other online payments such as by bank card.

Situational normality refers to the environment being perceived to conform to what is expected and to be conducive to a positive outcome. It has six sub-constructs. The first, general structural normality, refers to the prevalent conditions on the internet in relation to security and the degree to which expectations will be met. The degree to which the consumer considers the related institutions as competent, benevolent and acting with integrity reinforces situational normality. The degree of digital currency adoption can influence the consumer. This is related to the diffusion of innovation as consumers with different characteristics adopt an innovation depending on the degree and stage of the adoption of that innovation in general (Rogers, 1962). The last sub-construct, digital currency reputation from the consumer's perspective, can be influenced by a numbers of factors including reports in the media and word of mouth. The trajectory of a digital currency value in relation to other currencies, security breaches, successes or failures of organizations active in this area and developments in the regulatory framework can impact the reputation.

Structural assurance refers to the related regulations, laws and guarantees. Given the relatively immature area digital currencies exist in, these are both influential and changeable. From the consumers perspective there are similarities between the issues influencing situational normality and structural assurance but the former is more fluid and abstract while the latter is more rigid. This construct has seven sub-constructs. The first sub-construct focuses on the currency characteristics, not the technology. As digital currencies are an alternative to government backed currencies the structural assurance in both influences the level of trust in the former. The second sub-construct, the digital currency payment system, can be explored separately as it faces different challenges and competitors. The payment intermediary can fulfil the function of the payment but also act as a reputable independent third party reinforcing the consumers trust in a successful transaction. Digital currency peer-to-peer infrastructure can also be considered separately as it is the platform of the technology and fundamentally different to alternative currencies and payment systems. Self-imposed regulation is a factor as it is a way to mitigate the negative influence of the incomplete, immature and changeable current level of regulation.

The last sub-construct external regulation to the degree which it exists in different countries can influence the consumer.

4 Method

While the research this paper develops (Zarifis et al. 2014) was exploratory and led to proposing some extensions to a model of trust to cater for this context this paper attempts to validate that proposed model. The data was collected by questionnaires. The participants had made online purchases but had not necessarily used Bitcoin. It was considered that a sample from current Bitcoin users which would share characteristics of technology savvy early adopters would be less representative of consumers in general. The limitation here is the varying exposure to this immature technology that the participants would have. This is a point in the method where gauging the consumers' perspective was preferred over a more typical human computer interaction focused experimental approach. The questionnaire adapted scales from similar research on trust (McKnight et al. 2002; Wrightsman, 1991; Dobing et al. 1993) and the participant responses were recorded on a 7-point Likert scale. Participants were recruited online and were mostly from Europe and North America. The participants were volunteers and no payment was given.

The analysis was in three stages covering convergent, discriminant and nomological validity. These three parameters would give us construct validity (Bagozzi et al. 1991). The methods used were Principal Component Analysis, Confirmatory Factor Analysis and lastly Second-Order Models. Principle Component Analysis would show us whether the constructs were discriminant. As these construct had been explored in the qualitative phase in earlier research (Zarifis et al. 2014) there were strong indications of the validity of the model's constructs but as the previous sample was small the generalizability of the findings could not be claimed. Confirmatory Factor Analysis was implemented with Structural Equation Modeling to evaluate the convergent and discriminant validity of the sub-constructs. Lastly Second-Order Models were used to evaluate the relationships across constructs with LISREL 9.1.

5 Results

From 562 questionnaires received 528 were considered valid. The quantitative analysis of these questionnaires showed a degree of validity of the constructs that exceeded the minimum acceptable level. The hypothesized paths were significant. Additional links and adaptations of the model were explored but not found to be sufficiently supported. For the constructs that are relevant to all online transactions beyond supporting their validity in this context this research further supported their validity in online transactions by consumers in general. These previously validated constructs were personal innovativeness, disposition to trust, the sub-constructs of institutional trust that came from the general online trust model (McKnight et al. 2002), trusting beliefs, the general web experience and perceived site quality.

For the sub-constructs of institutional trust that are specific to the context of digital currencies this further supports the findings of the qualitative research which developed them (Zarifis et al. 2014) and suggests they are valid and generalizable. Digital currency adoption and digital currency reputation were particularly significant. Both of these are part of situational normality that appeared to have a stronger influence than structural assurance. This may be due to the early stage in the adoption of this technology and may subside when this technology matures and the uncertainty around certain issues is reduced. An alternative explanation of these results is that the sub-construct of situational normality is currently and will remain more significant for digital currencies due to the nature of the function they provide and the issues currencies and transactions involve.

| Construct | Sub-construct | Alpha |
|-------------------------|---|-------|
| Personal innovativeness | | 0.81 |
| Faith in humanity | Competence | 0.79 |
| | Benevolence | 0.86 |
| | Integrity | 0.83 |
| Trusting stance | | |
| Institution-based trust | Situational normality | |
| | -General, | 0.70 |
| | -Competence, Benevolence, Integrity | 0.78 |
| | -DC adoption | 0.86 |
| | -DC reputation | 0.85 |
| | Structural assurance | |
| | -Digital currency | 0.84 |
| | -Government backed currency | 0.85 |
| | -Payment intermediary | 0.81 |
| | -DC P2P infrastructure | 0.83 |
| | -Self-imposed regulation | 0.78 |
| | -External regulation | 0.81 |
| Trust in the retailer | Trusting beliefs | |
| | -Competence belief | 0.85 |
| | -Benevolence belief | 0.84 |
| | -Integrity belief | 0.83 |
| | Trusting intentions | |
| | -Willingness to depend | 0.81 |
| | -probability to follow advice | 0.84 |
| | -probability to give personal information | 0.83 |
| | -probability to make a purchase | 0.84 |
| General web experience | | 0.82 |
| Perceived site quality | | 0.86 |

Table 1. Construct and sub-construct reliability

6 Conclusion

Digital currencies appear to offer advantages such as convenient, immediate, low cost transactions and have attracted the interest from retailers and consumers (Giaglis and Kypriotaki, 2014). These currencies however face the challenge posed by trust that is particularly decisive in all online purchases and transactions (Jarvenpaa et al. 1998; Bhattacharjee, 2002; Zarifis and Kokkinaki, 2015). The main contribution of this research was to assess the validity of the digital currency trust model proposed (Zarifis et al. 2014). This is part of a thread of research into the potential value and limitations of digital cur-

rencies to the consumer (Giaglis and Kypriotaki, 2014). The data collected and analysed supported firstly the validity of the model and secondly that trust has related but distinct parts that should be given individual attention.

The consumer trust in digital currency mediated transactions model can benefit organizations utilizing digital currencies in a number of ways: Firstly those developing the technology that supports the currency itself but also organizations that want to use it can take the findings into consideration to reinforce trust. For example the disagreements over the future of Bitcoin (Hearn, 2015) reduce institutional trust significantly and should be avoided. Secondly organizations engaged in business to consumer e-commerce can adapt their models, marketing and outlets to build consumer trust more effectively. For example a retailer can compensate for the reduced institutional trust by increasing trust with further assurances and guarantees (Karimov and Brengman, 2014). Thirdly organizations engaged in business to consumer commerce faced with a range of payment systems and channels (Lazaris and Vrechopoulos, 2015; Zarifis and Kokkinaki, 2015) can make better informed choices about where digital currencies can fit into their multichannel strategy. The role of the payment intermediary who may offer the digital currency functionality also influences trust. The value of some steps that have already taken place are supported by the model such as some Bitcoin organizations volunteering themselves to be regulated so that institutional trust is enhanced (Gruber, 2013). Lastly the trust in digital currency enabled transactions model can provide clarity, a platform and structure for further research and discussion.

Further research is needed to confirm the findings with different samples, methodologies and digital currencies other than Bitcoin. Furthermore the constructs and sub-constructs may change in their significance and relevance over time. This is due to the volatile nature of this immature technology and its relationship with competing online payment systems which are also evolving.

References

- Akins, B.W., Chapman, J.L. and Gordon, J.M. (2013). "A Whole New World: Income Tax Considerations of the Bitcoin Economy", *Pittsburgh Tax Review*, URL: <http://ssrn.com/abstract=2394738> (visited on 25/07/2015).
- Bagozzi, R.P., Yi, Y. and Phillips L.W. (1991). "Assessing construct validity in organizational research". *Admin. Sci. Quart.* vol. 36, pp. 421.
- Bhattacharjee, A. (2002). "Individual trust in online firms: Scale development and initial test", *Journal of Management Information Systems*, vol. 19, iss. 1, pp. 211-241.
- Dobing, B. (1993). "Building trust in user-analyst relationships". Unpublished Doctoral Dissertation, Information and Decision Sciences Department, University of Minnesota, Minneapolis.
- Giaglis M.G. and Kypriotaki K.N. (2014). "Towards an Agenda for Information Systems Research on Digital Currencies and Bitcoin", *Business Information Systems*, vol. 183, pp 3-13.
- Gjermundrod, H. and Dionysiou, I. (2014). "Recirculating Lost Coins in Cryptocurrency Systems", *Business Information Systems*, vol. 183, pp 255-268.
- Gruber, S.M. (2013). "Trust, Identity, and Disclosure: Are Bitcoin Exchanges the Next Virtual Havens for Money Laundering and Tax Evasion?", *32 Quinnipiac L. Rev.*, pp.135, URL: <http://ssrn.com/abstract=2312110> (visited on 25/07/2015).
- Hayes, A. (2015). "Cryptocurrency value formation: an empirical analysis leading to a cost of production model for valuing bitcoin", *Ninth Annual Mediterranean Conference on Information System (MCIS)*, Samos, pending.
- Hearn, M. (2015). "Why is bitcoin forking? A tale of different visions", *Medium: Faith and Future*, URL: <https://medium.com/faith-and-future/why-is-bitcoin-forking-d647312d22c1> (visited on 25/07/2015).

- Hett, W. (2008). "Digital Currencies and the Financing of Terrorism", *Richmond Journal of Law & Technology*, vol. 15, iss. 2, URL: <http://law.richmond.edu/jolt/v15i2/article4.pdf> (visited on 25/07/2015).
- Internal Revenue Service: Notice 2014-21, (2014), URL: http://www.irs.gov/pub/irs-drop/n-14-21.pdf?utm_source=3.31.2014+Tax+Alert&utm_campaign=3.31.14+Tax+Alert&utm_medium=email (visited on 25/07/2015).
- Jarvenpaa, S., Knoll, K.D. and Leidner, D. (1998). "Is anybody out there? Antecedents of trust in global virtual teams". *Journal of Management Information Systems*, pp. 29-64.
- Karimov, F.P. and Brengman, M. (2014). "An examination of trust assurances adopted by top internet retailers: unveiling some critical determinants", *Electronic Commerce Research*, vol. 13, pp. 459-496.
- Lazaris, C. and Vrechopoulos, A.P. (2015). "The Interplay of Omniretailing & Store Atmosphere on Consumers' Purchase Intention towards the Physical Retail Store", 12th European, Mediterranean & Middle Eastern Conference on Information Systems (EMCIS), Athens, Greece.
- McKnight, D.H., Cummings, L.L. and Chervany, N.L. (1998). "Initial trust in new organizational relationships", *Academy of Management Review*, pp. 473-490.
- McKnight, D.H., Chaudhury V. and Kacmar C. (2002). "Developing and Validating Trust Measures for e-Commerce: An Integrative Typology", *Information Systems Research*, pp. 334-359.
- Nakamoto, S. (2009). "Bitcoin: A peer-to-peer electronic cash system", URL: <http://bitcoin.org/bitcoin.pdf> (visited on 25/07/2015).
- O'Brien, T. L. (2000). "Aided by Internet, identity theft soars", *The New York Times*, April 3.
- Pearce, W.B. (2007). "Trust in interpersonal relationships". *Speech monographs*, pp.832-835.
- Rogers, E. M. (1962). *Diffusion of innovations*. New York: Free Press.
- Sapuric, S. and Kokkinaki, A. (2014). "Bitcoin Is Volatile! Isn't that Right?". *Business Information Systems*, vol. 183, pp 255-268.
- Verbücheln, S. (2015). "How perfect offline wallets can still leak bitcoin private keys", Ninth Annual Mediterranean Conference on Information System (MCIS), Samos, pending.
- Wrightsman, L.S., (1991). "Interpersonal trust and attitudes toward human nature", In: Robinson, J.P., Shaver, P.R., Wrightsman, L.S. (Eds.), *Measures of Personality and Social Psychological Attitudes: Vol. 1: Measures of Social Psychological Attitudes*, Academic Press, San Diego, CA, pp. 373-412.
- Zarifis, A., Cheng, X., Efthymiou L. and Dimitriou, S. (2014). "Consumer Trust in Digital Currency Enabled Transactions", *Business Information Systems*, vol. 183, pp.241-254.
- Zarifis A. and Kokkinaki A. (2015). "The Relative Advantage of Collaborative Virtual Environments and Two-Dimensional Websites in Multichannel Retail", *Business Information Systems*, vol. 208, pp.233-244.

USING TIME-SERIES AND SENTIMENT ANALYSIS TO DETECT THE DETERMINANTS OF BITCOIN PRICES

Complete Research

Georgoula, Ifigenia, Athens University of Economics & Business, Athens, Greece, ifgeorgoula@aueb.gr

Pournarakis, Demitrios, Athens University of Economics & Business, Athens, Greece, pournadi@aueb.gr

Bilanakos, Christos, Athens University of Economics & Business, Athens, Greece, xmpilan@aueb.gr

Sotiropoulos, Dionysios, Athens University of Economics & Business, Athens, Greece, dsotirop@aueb.gr

Giaglis, M. George, Athens University of Economics & Business, Athens, Greece, giaglis@aueb.gr

Abstract

This paper uses time-series analysis to study the relationship between Bitcoin prices and fundamental economic variables, technological factors and measurements of collective mood derived from Twitter feeds. Sentiment analysis has been performed on a daily basis through the utilization of a state-of-the-art machine learning algorithm, namely Support Vector Machines (SVMs). A series of short-run regressions shows that the Twitter sentiment ratio is positively correlated with Bitcoin prices. The short-run analysis also reveals that the number of Wikipedia search queries (showing the degree of public interest in Bitcoins) and the hash rate (measuring the mining difficulty) have a positive effect on the price of Bitcoins. On the contrary, the value of Bitcoins is negatively affected by the exchange rate between the USD and the euro (which represents the general level of prices). A vector error-correction model is used to investigate the existence of long-term relationships between cointegrated variables. This kind of long-run analysis reveals that the Bitcoin price is positively associated with the number of Bitcoins in circulation (representing the total stock of money supply) and negatively associated with the Standard and Poor's 500 stock market index (which indicates the general state of the global economy).

Keywords: *Bitcoins; error correction; machine learning; sentiment analysis.*

1 Introduction

The Bitcoin is a digital currency which has recently emerged as a peer-to-peer payment system to facilitate transactions. It is not issued by any central bank or other financial institution but uses cryptographic methods and relies on an open-source software algorithm which verifies decentralized transactions and controls the creation of new Bitcoins. The large fluctuations of Bitcoin prices (especially within the year of 2013) and the huge increase in the capitalization of the associated market have given rise to a branch of the literature studying the factors which help to explain or predict the value of Bitcoins (Ciaian et al., 2014; Kristoufek, 2013; Li et al., 2014; Parker et al., 2012). In this paper, we study the dynamics governing the formation of Bitcoin prices by focusing on Twitter sentiment as an explanatory factor along with other economic and technological variables.

It has been argued in (Parker et al., 2012) that the price of Bitcoins is mainly driven by the interaction of supply and demand fundamentals (as it happens with other currencies or standard commodities). In this context, the impact of mining technology – which affects the production cost structure and thus the supply side of the market – on Bitcoin prices has been investigated by (Li et al., 2014). However, the supply of Bitcoins evolves according to a publicly known algorithm and the level of demand is not fully determined by the fundamentals of the underlying economy but also depends on expectations

about future price movements. Therefore, the standard economic theory might not adequately describe changes in Bitcoin prices and one should also take short-run speculative investment incentives or expectations into account. These expectations might be reflected in collective sentiment, thus raising the question of measuring public mood and studying its impact on the evolution of Bitcoin prices. In this context, (Kristoufek, 2013) uses search queries on Google trends and Wikipedia (representing the degree of public recognition or interest in Bitcoins) as a proxy for public sentiment and finds a positive correlation between these measures and the price of Bitcoins.

We extend the above line of reasoning here by constructing a sentiment ratio for Twitter users on a daily basis. Several measures of public mood associated with online social media have been suggested in the literature to predict the movement of stock market indexes (Bollen et al., 2011; Gilbert and Karahalios, 2010; Zhang et al., 2011). It has recently been argued that Twitter posts (related to Bitcoins) which express negative sentiments or uncertainty are negatively correlated with the price of Bitcoins (Kaminski and Gloor, 2014). In this paper, we perform sentiment analysis through the use of a state-of-the-art machine learning algorithm (namely Support Vector Machines). The econometric analysis of our time-series data implies that the Twitter sentiment ratio has a significantly positive impact on Bitcoin prices. The frequency of Wikipedia views and the level of mining difficulty (measured by the hash rate) are also positively associated but the exchange rate between the USD and the euro is negatively associated with the value of Bitcoins. For the cointegrated time series, the estimation of a vector error-correction model shows that the stock of Bitcoins has a positive long-run impact and the Standard & Poor's 500 index has a negative long-run impact on Bitcoin prices. Finally, we find that the price of Bitcoins adjusts to its long-run equilibrium value at a relatively high speed.

The rest of this paper proceeds as follows. Section 2 introduces the conceptual framework and describes the selected dataset. Section 3 suggests the methodology used to conduct the sentiment analysis and the set of econometric estimations. Section 4 derives the empirical results and discusses their implications. Section 5 concludes and provides directions for further research.

2 Theoretical Framework and Dataset

If we consider Bitcoin as a medium of exchange, then its price should be determined by standard supply and demand interactions (Ciaian et al., 2014; Parker et al., 2012). Fisher's (Fisher, 1922) equation of exchange associated with the quantity theory of money stipulates $MV=PT$ (where M is the nominal supply of money, V is the velocity of money circulation, P is the general price level and T is the size of the underlying economy). The nominal supply of Bitcoins is given by $M=P^B B$ (where P^B is the price of Bitcoins and B is the stock of Bitcoins in circulation), thus implying $P^B=PT/VB$. Therefore, the equilibrium price of Bitcoins (i.e. the price equalizing demand and supply) should be positively related to the general price level (P) and the size of the Bitcoin economy (T) but negatively related to the total stock of Bitcoins in circulation (B). In the same context, the mining difficulty can be used to measure the production cost of Bitcoins which affects the supply side of the market (Li et al., 2014). Furthermore, the level of demand for Bitcoins might be related to the general macroeconomic state of the global economy captured, for example, by alternative stock market indices (Ciaian et al., 2014). However, Bitcoins are also treated as an investment asset whose demand could be affected by speculative behavior associated with expectations and public feelings about their future price movements. These feelings might be captured by the degree of public recognition and interest in Bitcoins measured by the number of search queries in Google trends and Wikipedia (Kristoufek, 2013). Alternatively, public feelings can be measured through the sentiment analysis of posts related to Bitcoins on social media such as Twitter (Kaminski and Gloor, 2014).

On the grounds of this conceptual framework, we use time series data for eleven variables collected from 27 October 2014 to 12 January 2015 on a daily basis. All series have been transformed by taking natural logarithms to overcome the problems of many outliers and high skewness mainly associated with financial variables. The dependent variable is the price of Bitcoins given by the Bitstamp closing

price (*bcp*) in USD. There are four independent variables representing the supply and demand fundamentals of the market: First, the stock of Bitcoins in circulation (*totbc*) represents the total money supply. Second, the daily total number of unique transactions (*ntran*) describes the size of the Bitcoin economy. Third, the number of Bitcoin days destroyed for any given transaction (*bcdde*) measures the Bitcoin money velocity and is calculated by multiplying the number of Bitcoins in a transaction with the number of days elapsed since these coins were last spent. Fourth, the daily exchange rate (*exrate*) between the USD and the euro (\$/€) represents the price level of the global economy. We also include the Standard & Poor's 500 stock market index (*sp*) as an independent variable representing the general state of the global economy. The level of mining difficulty is captured by the hash rate (*hashrate*) measuring the processing power of the Bitcoin network. All these series were downloaded by *quandl.com*. We also use three proxies for the degree of public recognition and interest in Bitcoins: First, the number of Bitcoin searches in Wikipedia (*wikiviews*) downloaded from *bitcoinpulse.com*. Second, the (normalized) number of search queries in Google (*googleviews*) retrieved from Google Trends. Third, the daily number of Twitter posts (*ntweets*) related to Bitcoins as collected in our database. The last explanatory variable of our model is the daily sentiment ratio (*sent*) associated with Twitter posts. The methodology for constructing the time series of this sentiment ratio is described in the next section. The full set of variables included in the model is listed in Table 1 below.

| Name of variable | Description |
|------------------|--|
| <i>bcp</i> | Bitstamp daily closing price |
| <i>totbc</i> | Total daily number of Bitcoins in circulation |
| <i>ntran</i> | Total daily number of unique Bitcoin transactions |
| <i>bcdde</i> | Bitcoin days destroyed for any given transaction |
| <i>exrate</i> | Daily exchange rate between the USD and the euro (\$/€) |
| <i>sp</i> | Standard & Poor's 500 stock market daily index |
| <i>hash</i> | Processing power required for the secure operation of Bitcoin network (in billions of hashes per second) |
| <i>wiki</i> | Daily number of Bitcoin search queries on Wikipedia |
| <i>google</i> | Daily number of Bitcoin search queries on Google |
| <i>ntweets</i> | Daily number of Twitter posts related to Bitcoins |
| <i>sent</i> | Daily sentiment ratio of Twitter posts related to Bitcoins |

Table 1. The set of variables.

3 Methodology

3.1 Sentiment Analysis Methodology

We collected and analysed a set of over 2,125,243 tweets during a time period of 78 days, between October 27th 2014 and January 12th 2015. The data collection process focused on gathering tweets for the keywords "Bitcoin", "BTC" and "Bitcoins" along with their respective hashtags ["#Bitcoin", "#BTC", "#Bitcoins"]. This task was accomplished by parsing the official streaming API of Twitter with Python and MySQL for storing the data. Appropriate wrappers were deployed in our dedicated Ubuntu server to ensure the process against runtime errors during network downtime. The resulting dataset was subsequently submitted to a series of data clearing and pre-processing operations. The data preparation process involved text tokenization into words, elimination of English stop-words and words with less than three characters, and stem extraction from each word. Therefore, the final version

of our corpus was formed by a collection of purified documents where each document contained the text from a single tweet.

3.1.1 Corpus Vectorization

A natural approach towards sentiment analysis is through a mathematical representation of the corpus via the employment of the standard Vector Space Model (VSM) originally introduced in (Salton et al., 1975). The main idea behind VSM is to transform each document d into a vector containing only the words that belong to the document and their frequency by using the “bag of words” representation. According to VSM, each document is represented exclusively by the words it contains by tokenizing sentences into elementary term (word) elements losing the associated punctuation, order and grammar information. The underlying mathematical abstraction imposed by VSM entails a mapping which transforms the original purified document to its corresponding bag of terms representation. This transformation can be formulated by the following equation:

$$\varphi: d \rightarrow \varphi(d) = [tf(t_1, d), tf(t_2, d), \dots, tf(t_M, d)] \in \mathbb{R}^M \quad (1)$$

where $tf(t_i, d_j)$ is the normalized frequency of the term t_i in document d_j given by:

$$tf(t_i, d_j) = f(t_i, d_j) / \max\{f(t, d_j) : t \in d_j\} \quad (2)$$

where $f(t_i, d_j)$ is the absolute frequency of the term t_i in document d_j . Based on the adopted mathematical formulation for the fundamental notions of corpus and dictionary, a corpus D of n documents and a dictionary T of M terms may be represented by $D = \{d_1, d_2, \dots, d_n\}$ and $T = \{T_1, T_2, \dots, T_M\}$. Having in mind Eq.1 and the formal definitions of corpus and dictionary, the mathematical representation for corpus in the context of VSM can be done through the document-term matrix:

$$D = \begin{bmatrix} tf(t_1, d_1) & \dots & tf(t_M, d_1) \\ \vdots & \ddots & \vdots \\ tf(t_1, d_n) & \dots & tf(t_M, d_n) \end{bmatrix} \quad (3)$$

where N is typically quite large, resulting in a sparse VSM representation such that a few matrix entries are non-zero. In order to mitigate the effect related to the complete loss of context information around a term, we incorporated the term-frequency inverse document frequency (tf-idf) weighting scheme according to which each term t_i is assigned a weight of the form:

$$w_i = idf(t_i, D) = \log(|D| / |\{d \in D : t_i \in d\}|) \quad (4)$$

so that the relative importance of each term for the given corpus is taken into consideration.

3.1.2 Support Vector Machines

Sentiment analysis was conducted through the utilization of a state-of-the-art classifier, namely Support Vector Machines (SVMs). SVMs are non-linear classifiers that were initially formulated by (Vapnik, 2013), operating in higher-dimensional vector spaces than the original feature space of the given dataset. Letting $S = \{(\vec{x}_i, y_i) \in \mathbb{R}^n \times \{-1, +1\}, \forall i \in [m]\}$ be the set of m training patterns with associated binary labels, such that -1 denotes the class of negative sentiment and $+1$ the class of positive sentiment, the learning phase of the SVMs involved solving the following quadratic optimization problem:

$$\min_{\vec{w}, \xi, b} \frac{1}{2} \|\vec{w}\|^2 + C \sum_{i=1}^m \xi_i$$

$$s.t. y_i (\langle \vec{w}, \vec{x}_i \rangle + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad \forall i \in [m]$$

The previous primal optimization problem has a corresponding dual one which gives rise to a discrimination function of the form:

$$g(\vec{x}) = \sum_{i \in SV} \alpha_i^* y_i \langle \vec{x}, \vec{x}_i \rangle + b^* \tag{5}$$

where $\{\alpha_i^*, i \in [m]\}$ and b^* denote the optimal solutions for the corresponding optimization variables and SV is the subset of training patterns associated with positive Lagrange multipliers. Given that the training patterns appear only in dot product terms of the form $\langle \vec{x}_i, \vec{x} \rangle$, a positive definite kernel function such as $K(\vec{u}, \vec{v}) = \Phi(\vec{u}) \Phi(\vec{v})$ can be employed in order to implicitly map the input feature space into a higher-dimensional vector space and compute the dot product. In this paper, we utilized the Gaussian kernel function defined by the following equation.

$$K(\vec{x}, \vec{y}) = \exp(-\|\vec{x} - \vec{y}\|^2 / 2\sigma^2) \tag{6}$$

3.1.3 Sentiment Classification

Our sentiment classification process was further divided into the corresponding training and testing stages. The training stage is an essential part of our methodology, since the application of SVMs on such a large amount of tweets requires a reasonable amount of labelled data (i.e. tweets already classified as positive, negative or neutral, based on a business perspective classification). This ensures that the SVM algorithm runs with accuracy, providing robust results that limit the amount of fault. These labelled data are in turn utilized by the SVM algorithm as a benchmark to score the number of tweets that are in scope of the sentiment exercise. In order to create a reasonable amount of labelled data, we manually labelled a set of collected tweets in terms of sentiment as positive (1), neutral (0) or negative (-1). The testing stage, on the contrary, aims at testing the accuracy and validity of the SVM algorithm on the largest subset of the dataset that was not previously classified.

In order to demonstrate the validity of the SVM algorithm for the sentiment classification problem, we adopted the standard 10-fold cross validation process on the previously labelled Tweets and measured the corresponding training and testing sentiment classification accuracy. Each fold involved splitting the complete set of pre-labelled samples into a 95% training data - 5% testing data ratio, where the first subset of data instances was used to build the classifier and the latter for assessing its ability to infer the sentiment polarity of unseen data patterns. The scores of the SVM classifier are summarized in Table 2. The sentiment categorization for the rest of the unlabelled data patterns was conducted by exploiting the complete set of pre-labelled data instances so that the trained classifier accumulated the maximum amount of available knowledge for the problem of sentiment classification.

| |
|--|
| Performing 10-fold cross-validation on labeled data on 2 classes: (346L, 800L) (48L, 800L) (1L, 346L) (1L, 48L) (394L,) |
| Accuracy per fold: [0.9 0.875 0.9 1. 0.875 0.875 0.87179487 0.87179487 0.89473684 0.89473684] |
| Mean Accuracy: 0.896 (+/- 0.073) |
| Precision per fold: [0.8974359 0.875 0.8974359 1. 0.875 0.875 0.87179487 0.87179487 0.89473684 0.89473684] |
| Mean Precision: 0.895 (+/- 0.073) |

| |
|---|
| Recall per fold: [1. 1. 1. 1. 1. 1. 1. 1. 1. 1.] |
| Mean Recall: 1.000 (+/- 0.000) |
| F-Score per fold: [0.94594595 0.93333333 0.94594595 1. 0.93333333 0.93333333 0.93150685 0.93150685 0.94444444 0.94444444] |
| Mean F-Score: 0.944 (+/- 0.039) |

Table 2. Sentiment Classification Scores.

3.2 Econometric Methodology

The econometric approach follows a number of steps associated with the analysis of time series data. The first step requires investigating whether the variables are stationary or not. A non-stationary series is integrated of order d (I(d)) if it becomes stationary by taking its differences d times. If we include a non-stationary series in a regression model, the ordinary least squares (OLS) estimators are not consistent and the standard statistical tests are not valid. Therefore, we might infer a statistically significant causal relationship between a pair of variables although such a relationship does not exist (Engle and Granger, 1987). In order to avoid this problem of spurious regression, we conduct several stationarity tests for each series y_t . In particular, we start with the augmented Dickey-Fuller (ADF) test based on an autoregressive model of order one:

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t \tag{7}$$

If $a_1=1$, the series has a unit root (i.e. it is non-stationary). We subtract y_{t-1} from both sides of (1) and define $\theta = a_1 - 1$, $\Delta y_t = y_t - y_{t-1}$ to get:

$$\Delta y_t = a_0 + \theta y_{t-1} + \varepsilon_t \tag{8}$$

If the series has a trend, we must explicitly include time as an explanatory variable in (7). Furthermore, we can add p lags of Δy_t to account for the dynamics of the process:

$$\Delta y_t = a_0 + \delta t + \theta y_{t-1} + \sum_{j=1}^p \gamma_j \Delta y_{t-j} + \varepsilon_t \tag{9}$$

The ADF test has a null hypothesis of a unit root ($\theta=1$) against the alternative of stationarity. In order to strengthen the validity of inferences, we also use the Phillips-Perron (PP) unit root test (Phillips and Perron, 1988). This can be viewed as a Dickey-Fuller test which has been made robust to serial correlation by using the Newey-West heteroscedasticity and autocorrelation-consistent (HAC) standard errors. Finally, we verify our results by also applying the KPSS test which has a null hypothesis of stationarity against the alternative of a unit root (Kwiatkowski et al., 1992).

A non-stationary series which is integrated of order d can be made stationary by taking its differences d times. However, the use of differences in a regression model does not allow determining potential long-run relationships between the variables. Therefore, we would like to use the levels of variables but also avoid the problem of spurious regression. If the variables are cointegrated then a regression model involving their levels yields consistent OLS estimators (Engle and Granger, 1987). The series $y_t, x_{1t}, \dots, x_{kt}$ are cointegrated CI(d,b) if all of them are integrated of order d and there exists a linear combination of these series which is integrated of order d-b. The standard example involves d=b=1, implying that the series are I(1) and there exists a linear combination $u_t = y_t - a_0 - a_1 x_{1t} - \dots - a_k x_{kt}$ which is stationary. Then, there is a long-run relationship between the cointegrated variables given by:

$$u_t = y_t - a_0 - a_1 x_{1t} - \dots - a_k x_{kt} + u_t \tag{10}$$

The vector $[1, -\alpha_0, -\alpha_1, \dots, -\alpha_k]$ is the cointegrating vector and might not be unique for the case of multiple (more than two) variables. The lagged series u_{t-1} is the error correction term measuring deviations from the long-run equilibrium. In this context, the second step of our econometric analysis involves testing for cointegration between the series which have the same order of cointegration. For the multivariate case, the cointegration test is based on Johansen's method calculating a trace statistic to specify the number of cointegrating vectors (Johansen, 1995).

After conducting the cointegration tests, we proceed with two separate sets of estimations. On the one hand, we rely on several OLS regressions to identify short-run relationships between the price of Bitcoins and the set of independent variables. Of course, these regressions require the transformation of non-stationary series (by taking their differences) to render them stationary. On the other hand, for the series which are found to be cointegrated we build a vector error-correction model (VECM) to detect the existence of long-run relationships (Enders, 2008). For ease of exposition, let us focus here on the case of two series x_t and y_t . If these series are cointegrated, there exist unique values of α_0 and α_1 such that $u_t = y_t - \alpha_0 - \alpha_1 x_t$ is stationary. If we think of y_t as the dependent variable and x_t as an exogenous regressor, the single-equation error-correction model is written as:

$$\Delta y_t = \beta_0 + \beta_1 \Delta x_t + \lambda u_{t-1} + \varepsilon_t = \beta_0 + \beta_1 \Delta x_t + \lambda (y_{t-1} - \alpha_0 - \alpha_1 x_{t-1}) + \varepsilon_t \tag{11}$$

The VECM extends (10) by allowing the joint evolution of x_t and y_t and by putting p lags on the right-hand side of both equations involved in the associated system:

$$\Delta y_t = \gamma_1 + \sum_{i=1}^p \delta_{1i} \Delta y_{t-i} + \sum_{i=1}^p \theta_{1i} \Delta x_{t-i} + \lambda_1 (y_{t-1} - \alpha_0 - \alpha_1 x_{t-1}) + \varepsilon_{1t} \tag{12}$$

$$\Delta x_t = \gamma_2 + \sum_{i=1}^p \delta_{2i} \Delta y_{t-i} + \sum_{i=1}^p \theta_{2i} \Delta x_{t-i} + \lambda_2 (y_{t-1} - \alpha_0 - \alpha_1 x_{t-1}) + \varepsilon_{2t} \tag{13}$$

Since the terms Δy_{t-i} , Δx_{t-i} and u_{t-1} are stationary, the OLS estimators of equations (12) and (13) are consistent. The coefficients δ_{ji} and θ_{ji} represent the short-run dynamics whereas α_0 and α_1 describe the long-run relationship between x_t and y_t . The parameters λ_1 and λ_2 contain information on the speed of adjustment to the long-run equilibrium by showing the correction of the previous period's disequilibrium error taking place in period t . The next section applies the above methodology and discusses the empirical results concerning the short-run and long-run determinants of Bitcoin prices.

4 Results

4.1 Stationarity and Cointegration

We mainly rely on the statistical program STATA but also use R and Matlab where necessary to verify our empirical results. First, we study the stationarity of each series by using the ADF, the PP and the KPSS tests. Since the ADF test yields ambiguous results concerning the stationarity of some independent variables, we turn to the PP unit root test which uses the Newey-West HAC standard errors to account for serial correlation. The results are summarized in Table 3 which shows the PP test statistic and the associated p-value for all series as well as for the first differences of non-stationary series. We conclude that the variables logbcp, logtotbc, logsp and logexrate are I(1), whereas for all other variables the null hypothesis of a unit root is rejected. These results are verified by the KPSS test having the opposite null hypothesis of stationarity.

| Variable | PP-test statistic | p-value |
|----------|-------------------|---------|
| logbc | -1.656 | >0.1 |
| d.logbc | -8.171 | <0.01 |

| | | |
|-------------|--------|-------|
| logntran | -4.460 | <0.01 |
| logtotobc | -2.057 | >0.1 |
| d.logtotbc | -6.240 | <0.01 |
| logbcdde | -6.406 | <0.01 |
| logwiki | -7.995 | <0.01 |
| loggoogle | -6.096 | <0.01 |
| logntweets | -6.961 | <0.01 |
| Logsp | -2.875 | >0.1 |
| d.logsp | -7.213 | <0.01 |
| loghash | -8.615 | <0.01 |
| logexrate | -1.786 | >0.1 |
| d.logexrate | -9.499 | <0.01 |
| logsent | -6.635 | <0.01 |

Table 3. Phillips-Perron (PP) unit root tests.

We proceed with the cointegration analysis to check whether there exists a long-run relationship between the four non-stationary series. For this purpose, we use Johansen's trace test for multivariate cointegration. The procedure starts with testing for zero cointegrating vectors and then accepts the first null hypothesis that is not rejected. As shown in Table 4, we reject the null hypothesis of no cointegration but fail to reject the null hypothesis of at most one cointegrating vector. Therefore, we accept that there is one cointegrating equation in the model.

| Cointegrating vector | Trace statistic | 5% critical value |
|----------------------|-----------------|-------------------|
| 0 | 53.3435 | 47.21 |
| 1 | 24.9293* | 29.68 |
| 2 | 10.7067 | 15.41 |
| 3 | 1.7893 | 3.76 |

Table 4. Johansen's test for cointegration.

The short-run impact of independent variables on the price of Bitcoins is studied through the estimation of several OLS regression models in the next subsection. The long-run relationship between the cointegrated series is determined by using a vector error-correction model (VECM).

4.2 OLS Estimates

All the regressions run in this section use the Newey-West HAC standard errors. Non-stationary series have been transformed by taking their first differences to guarantee non-spurious results. We first consider a regression containing the full set of explanatory variables. Table 5 shows that the variables logwiki (representing the degree of public recognition or interest in Bitcoins) and loghash (measuring the mining difficulty) have a positive impact on Bitcoin price, whereas the effect of the exchange rate between the USD and the euro (reflecting the price level of the global economy) is negative. Even more interestingly, the sentiment ratio of Twitter users positively affects the price of Bitcoins. Since the removal of non-significant factors increases the overall F-statistic of the model, we keep only the significant variables and add lags in the regression. Table 6 shows the results with three lags for each

explanatory variable. Apart from own-price effects, we note that the second lag of Wikipedia views and the second lag of the Twitter sentiment index also have a significantly positive effect on current Bitcoin prices. The list of variables correlated with the price of Bitcoins in the short run is shown in Table 7.

Regression with Newey-west standard errors
maximum lag: 5

Number of obs = 76
F(10, 65) = 5.74
Prob > F = 0.0000

| D. logbcp | Coef. | Newey-west Std. Err. | t | P> t | [95% Conf. Interval] |
|------------------|-----------------------|----------------------|---------------|----------------|---|
| logntran | -.0009874 | .0378949 | -0.03 | 0.979 | -.0766687 .0746939 |
| logtotbc D1. | -137.6733 | 148.6298 | -0.93 | 0.358 | -434.5074 159.1608 |
| logbcdde | -.005857 | .0042119 | -1.39 | 0.169 | -.0142688 .0025548 |
| logwiki | .0202568 | .007452 | 2.72 | 0.008 | .0053742 .0351394 |
| loggoogle | -.0237488 | .0317945 | -0.75 | 0.458 | -.0872467 .0397491 |
| logtweets | .000973 | .022617 | 0.04 | 0.966 | -.0441963 .0461423 |
| logsp D1. | -.1239438 | .6570604 | -0.19 | 0.851 | -1.436184 1.188296 |
| loghash | .0072955 | .002659 | 2.74 | 0.008 | .001985 .0126059 |
| logexrate D1. | -1.935874 | 1.053349 | -1.84 | 0.071 | -4.039557 .1678083 |
| logsent _cons | .6890404 -.0670151 | .2568672 .4132169 | 2.68 -0.16 | 0.009 0.872 | .1760412 -.8922663 1.20204 .7582361 |

Table 5. OLS regression with the full set of explanatory variables.

Regression with Newey-west standard errors
maximum lag: 5

Number of obs = 73
F(19, 53) = 10.37
Prob > F = 0.0000

| D. logbcp | Coef. | Newey-west Std. Err. | t | P> t | [95% Conf. Interval] |
|------------------|-----------|----------------------|-------|-------|----------------------|
| logbcp LD. | .2486992 | .0907913 | 2.74 | 0.008 | .066595 .4308033 |
| L2D. | -.2584576 | .0808432 | -3.20 | 0.002 | -.4206085 -.0963067 |
| L3D. | -.0360947 | .1621652 | -0.22 | 0.825 | -.3613569 .2891675 |
| logwiki --- | .0197162 | .0068254 | 2.89 | 0.006 | .0060263 .0334062 |
| L1. | -.000763 | .0088237 | -0.09 | 0.931 | -.0184612 .0169351 |
| L2. | .017811 | .008851 | 2.01 | 0.049 | .000058 .0355639 |
| L3. | -.0001393 | .0083484 | -0.02 | 0.987 | -.0168841 .0166054 |
| loghash --- | .0088932 | .0034975 | 2.54 | 0.014 | .0018782 .0159082 |
| L1. | .0071992 | .0058986 | 1.22 | 0.228 | -.0046319 .0190304 |
| L2. | -.0105373 | .0044826 | -2.35 | 0.022 | -.0195282 -.0015464 |
| L3. | .001057 | .0046555 | 0.23 | 0.821 | -.0082807 .0103947 |
| logexrate D1. | -1.674214 | .7754822 | -2.16 | 0.035 | -3.229635 -.1187941 |
| LD. | 1.54933 | 1.329604 | 1.17 | 0.249 | -1.117517 4.216177 |
| L2D. | -.1552209 | .7811126 | -0.20 | 0.843 | -1.721934 1.411492 |
| L3D. | -.9975245 | .8014653 | -1.24 | 0.219 | -2.60506 .6100113 |
| logsent --- | .8069225 | .283992 | 2.84 | 0.006 | .2373067 1.376538 |
| L1. | -.5240742 | .3189217 | -1.64 | 0.106 | -1.16375 .1156016 |
| L2. | .8888849 | .5158606 | 1.72 | 0.091 | -.1458004 1.92357 |
| L3. | -.2830255 | .3074419 | -0.92 | 0.361 | -.8996759 .3336249 |
| _cons | -.3719047 | .2401851 | -1.55 | 0.127 | -.8536549 .1098455 |

Table 6. OLS regression with three lags for each significant explanatory variable.

| Variable | Effect on Bitcoin price |
|-----------------------|-------------------------|
| Wikipedia views | Positive |
| Hash rate | Positive |
| Sentiment ratio | Positive |
| USD/EUR exchange rate | Negative |

Table 7. Short-run influencers of Bitcoin prices.

4.3 Vector Error-Correction Model

The long-run relationship between the cointegrated variables is now determined through a VECM with four lags (based on the information criteria). The first part of Table 8 shows the short-run dynamics as well as the speed of adjustment to the long-run equilibrium. The second lag of *logsp* has a negative impact and the lagged difference of *logexrate* is now found to have a positive impact on the price of Bitcoins. It can be seen that 31.25% of the gap between the Bitcoin price in period t-1 and its equilibrium value tends to be reversed in period t. In other words, if the Bitcoin price is too high then it falls back towards its equilibrium level relatively quickly. The second part of Table 8 shows the coefficients of the cointegrating equation. An increase in the stock of Bitcoins leads to an increase in the Bitcoin price (contrary to our expectations), while an increase in the Standard and Poor’s 500 stock market index (showing an improvement in the state of the global economy) negatively affects the price of Bitcoins in the long run. The last result potentially reflects the fact that investment in stocks and investment in Bitcoins are treated as substitutes. Table 9 shows the set of variables having a long-run effect on Bitcoin prices.

```

-----+-----
                Coef.  Std. Err.  z    P>|z|    [95% Conf. Interval]
-----+-----
d_logbcp
  _ce1
  L1.      -0.3125402   0.1048414   -2.98  0.003   -0.5180255   -0.1070549

  logbcp
  LD.       0.3203674   0.1287212    2.49  0.013    0.0680785    0.5726564
  L2D.     -0.208652    0.1417118   -1.47  0.141   -0.486402    0.069098
  L3D.     -0.0429716   0.137881    -0.31  0.755   -0.3132135   0.2272702
  L4D.     -0.1119016   0.1336133   -0.84  0.402   -0.373779    0.1499757

  logtotbc
  LD.     -0.2269944   0.2396149   -0.95  0.343   -0.696631    0.2426423
  L2D.     -0.1192837   0.213636    -0.56  0.577   -0.5380026   0.2994353
  L3D.     -0.2355137   0.2049175   -1.15  0.250   -0.6371446   0.1661172
  L4D.     -0.1890374   0.2021341   -0.94  0.343   -0.6000000   0.1661172

  logsp
  LD.       0.6393147   0.8691665    0.74  0.462   -1.06422    2.34285
  L2D.     -1.832229    0.913669   -2.01  0.045   -3.622988   -0.0414707
  L3D.     -0.8660657   0.9259208   -0.94  0.350   -2.680837    0.9487058
  L4D.     -0.5406169   0.8668753   -0.62  0.533   -2.239661    1.158427

  logexrate
  LD.       3.698214    1.287785    2.87  0.004    1.174203    6.222226
  L2D.     -0.5359462   1.395253   -0.38  0.701   -3.270591    2.198699
  L3D.       0.177157    1.337561    0.13  0.895   -2.444415    2.798729
  L4D.     -0.0301283   1.141858   -0.27  0.789   -2.268128    2.207872

  _cons     0.0005305   0.0728064    0.01  0.994   -0.1421673   0.1432284
-----+-----

Cointegrating equations
Equation      Parms   chi2    P>chi2
-----+-----
  _ce1                3    88.02165  0.0000

Identification:  beta is exactly identified

Johansen normalization restriction imposed
-----+-----
beta |      Coef.  Std. Err.  z    P>|z|    [95% Conf. Interval]
-----+-----
  _ce1 |
  logbcp |          1
  logtotbc | 15.9535   4.122128    3.87  0.000    7.874277    24.03272
  logsp | -4.210831  0.7492988  -5.62  0.000   -5.67943    -2.742233
  logexrate | -5.230419  1.434541  -3.65  0.000   -8.04691    -2.41393
  _cons | -236.1674
-----+-----

```

Table 8. Vector error-correction model with four lags.

| Variable | Effect on Bitcoin price |
|--------------------|-------------------------|
| Number of Bitcoins | Positive |
| S&P's 500 index | Negative |

Table 9. Long-run influencers of Bitcoin prices.

5 Conclusion

This paper tried to shed light on the factors determining the price of Bitcoins in the short-run as well as in the long-run. We built an empirical model incorporating multiple economic and technological variables but also extended the existing literature by taking Twitter sentiment into account. Specifically, we used a state-of-the-art machine learning algorithm (Support Vector Machines) to measure the sentiment ratio of Twitter users concerning Bitcoins on a daily basis. After dealing with issues of stationarity and cointegration, we estimated several regression models indicating that our Twitter sentiment ratio has a positive short-run impact on Bitcoin prices. In other words, evidence shows that measurements of collective mood based on the appropriate sentiment analysis can help to predict short-run movements in the value of Bitcoins. Furthermore, the price of Bitcoins has been found to be positively affected by the number of Wikipedia search queries. This implies that a higher degree of public recognition or interest in Bitcoins increases their market price. Similarly, an increase in the hash rate has a positive effect on Bitcoin prices. This is hardly surprising, since the hash rate indicates the mining difficulty or marginal production cost of Bitcoins and thus normally exerts an upward pressure on their price. On the contrary, our estimations revealed a negative short-run relationship between the price of Bitcoins and the exchange rate between the USD and the euro. To the extent that this exchange rate represents the general level of prices, its inverse relationship with the value of Bitcoins contrasts the prediction of Fisher's equation associated with the quantity theory of money. For the set of cointegrated variables, we estimated a VECM to identify the underlying long-run relationships. The analysis revealed that the stock of Bitcoins has a positive long-run impact on their price. This is also a counter-intuitive result, since the number of Bitcoins in circulation measures the total supply of money which would be expected to have a negative effect on Bitcoin prices. The Standard and Poor's 500 index was found to have a negative impact on Bitcoin prices in the long run, implying that stocks and Bitcoins are treated as substitutes by investors. More specifically, a decrease in the Standard and Poor's 500 index induces investors to sell their stocks and substitute them for Bitcoins. Finally, the speed at which the price of Bitcoins adjusts to its long-run equilibrium value is relatively high. In particular, about one half of the deviation between the current and the equilibrium level of Bitcoin prices is already corrected within the next period.

The empirical model developed above can be improved and extended in multiple ways. First of all, a larger dataset should be used to check whether the conclusions reached here remain valid or not. Second, a vector autoregressive (VAR) model might be used (instead of simple OLS regressions) to study the short-run dynamics of price formation. Finally, alternative sentiment indices for social media users might be constructed (by applying the appropriate algorithmic processes) to explore their short-run and long-run effectiveness in explaining the price of Bitcoins. These extensions are left for future research.

Acknowledgments. This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) – Research Funding Programs: Thalys – Athens University of Economics and Business – Software Engineering Research Platform; Aristeia – Athens University of Economics and Business – Herding Behavior and Asymptotic Learning in Electronic Social Media – Sociomine.

References

- Bollen, J., Mao, H., Zeng, X., 2011. Twitter mood predicts the stock market. *Journal of Computational Science* 2, 1–8.
- Ciaian, P., Rajcaniova, M., Kancs, d'Artis, 2014. The Economics of BitCoin Price Formation. arXiv preprint arXiv:1405.4498.
- Enders, W., 2008. *Applied econometric time series*. John Wiley & Sons.
- Engle, R.F., Granger, C.W., 1987. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society* 251–276.
- Fisher, I., 1922. Purchasing power of money: Its determination and relation to credit interest and crises, *Rev.*
- Gilbert, E., Karahalios, K., 2010. Widespread Worry and the Stock Market., in: *ICWSM*. pp. 59–65.
- Johansen, S., 1995. *Likelihood-based inference in cointegrated vector autoregressive models*. OUP Catalogue.
- Kaminski, J., Gloor, P., 2014. Nowcasting the Bitcoin Market with Twitter Signals. arXiv preprint arXiv:1406.7577.
- Kristoufek, L., 2013. BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific reports* 3.
- Kwiatkowski, D., Phillips, P.C., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics* 54, 159–178.
- Li, X., Wang, C.A., Wang, Q., 2014. Exploring the Determinants of Bitcoin Exchange Rate. Qi, Exploring the Determinants of Bitcoin Exchange Rate (October 15, 2014).
- Parker, J., Buchholz, M., Warren, J., Delaney, J., 2012. Buchholz \$\$\$\backslash\$ &\$} al.-Bits and Bets Information, Price Volatility, and Demand for Bitcoin (Independent).
- Phillips, P.C., Perron, P., 1988. Testing for a unit root in time series regression. *Biometrika* 75, 335–346.
- Salton, G., Wong, A., Yang, C.-S., 1975. A vector space model for automatic indexing. *Communications of the ACM* 18, 613–620.
- Vapnik, V., 2013. *The nature of statistical learning theory*. Springer Science & Business Media.
- Zhang, X., Fuehres, H., Gloor, P.A., 2011. Predicting stock market indicators through twitter “I hope it is not as bad as I fear.” *Procedia-Social and Behavioral Sciences* 26, 55–62.

EXTRACTING GREEK ELECTIONS TWEET'S CHARACTERISTICS

Completed Research

Kalaidopoulou, Katerina¹, Department of Management Science and Technology, Athens University of Economics and Business, Athens, Greece, kkalaid@gmail.com

Romanou, Angeliki¹, Department of Management Science and Technology, Athens University of Economics and Business, Athens, Greece, agromanou@gmail.com

Lekakos George¹, Department of Management Science and Technology, Athens University of Economics and Business, Athens, Greece, glekakos@aueb.gr

Abstract

Social media offer platforms that anyone can use, giving the opportunity to share information among networks in an easy and interactive way. It is not a surprise that social media marketing has become a primary focus on both digital and traditional revenue models of businesses. In this work, information sharing by users in the context of Twitter is studied, by modeling message's characteristics and users' behavior about Greek 2015 January elections. A detailed data set about tweets' characteristics such as length, existence of URLs or hashtags and mentioning of other users, is collected after the elections day, and the relationships between related users and network's responses on the shared tweets, are examined. An unsupervised clustering model is implemented on tweets' characteristics using CRISP-DM methodology. The empirical results suggest the existence of different content groups, such as tweets with extensive text, URLs and hashtags which can be characterized as "Linked" type of shared content.

Keywords: Machine learning, social media analysis, tweet clustering, Twitter

1. Introduction

Social media have become the most popular way to create, share and exchange information, pictures, music, videos, thoughts in digital communities. Facebook, Twitter, YouTube, LinkedIn are just a few of them. As people follow, like, share, tweet, retweet, tag, rate and text one another, they become part of an enormous social network, providing the opportunity to be extracted and analyzed towards identifying users' behavioral patterns and performing more effective information sharing and diffusion. The identification of the critical characteristics of the messages, that enables the maximization of its impact in a network, is a critical business goal and important users' need which paves the way for extensive research from both industry and academic community.

This need is increasing when it comes to analyze collected information under government elections' context. Related studies on this matter have been conducted on many different

¹ ELTRUN: The e-Business Research Center

topics such as elections, crises, televised events (Bruns and Stieglitz 2012). According to social sciences' studies, user has a great influence on the message's impact on the network. However, while the message, i.e. tweet, itself may influence the response it will get, tweet's characteristics, such as length tend to be neglected when mining them. In this work we focused on using a content-based approach on grouping messages/tweets on a time frame of a week after Greek January elections which are considered to be one of the most critical and diverged elections for Greece in the last fifty years. Extracting groups of similar tweets based on their characteristics (length, existence of URLs etc.) can lead to better understanding of the different types of tweets, correlating each of their characteristics with the network's response.

We contribute to this field by exploring Twitter social network so as to identify different types of tweets based on their characteristics such as length, hashtags etc. Which types of posted tweets maximize the network's response? How can we predict which characteristics are most related to high response and alternatively indicate high impact on the network? This research aims to approach these questions and pave the way for extensive future work on content characteristic-based analysis. We also seek to introduce an alternative way of tweets' segmentation, based on their characteristics and correlate them with users' attributes and network's response.

2. Research background

Previous studies have explored how social media users post, reply or forward messages, i.e. in Twitter how they tweet, retweet and reply. A more specific study conducted by Boyd et al. (2010) analyzed how, why and what people retweet in Twitter network, concluding that Twitter is mainly seen as a conversational environment. Though, in the context of user-centric classification, previous research has indicated that weak ties (in the form of unidirectional links) are more likely to engage in the social exchange process of content sharing. (Shi et al, 2013). Moreover, Pennacchiotti, Marco, and Ana-Maria Popescu (2011) have attempted to classify users based on a comprehensive set of features derived from such user information. Additionally to this research, Stieglitz, Stefan, and Linh Dang-Xuan (2013) seek to explain "whether sentiment occurring in social media content is associated with a user's information sharing behavior" (2013).

Many scholars have also studied Twitter activity in the context of individual case studies, which represent many different topics, such as elections, crises, televised events etc. A more generic study, conducted by Bruns and Stieglitz (2012) implies that "standard response to the emergence of breaking news and other acute events is the tendency to find, share, and re-share relevant information, resulting in a high rate of URLs and retweets". On the other hand, in live and streaming events "Twitter acts as a backchannel, containing mainly original commentary that does not engage with the tweets of others or provide a substantial number of links to further information".

As a result, it is indicated that in different types of tweets, different characteristics may drive awareness and response from network's peers. Focusing more on elections, we found that

other researches cover respective events from other countries in terms of predicting the elections' results (Tumasjan et al, 2010) or analyzing public communication (Bruns and Burgess, 2011). In our case we try to identify the most important characteristics of the message that maximize the response it will get from the network. Literature has issued the need for further content-based analysis on social media published material. However, none of the so-far work has approached content analysis on the context of Greek 2015 controversial elections. Thus, we contribute to this field by performing an exploratory analysis based on tweets' content characteristics, deriving distinct clusters of tweets based on tweets' and their publishers' attributes. We are using the CRISP-DM methodology (Chapman et al, 2000), a concrete methodology for unsupervised machine learning, whose phases are analyzed below.

3. Methodology and Data

Our analysis process was implemented according to the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology for approaching data mining problems. In order for the methodology to meet this work's goals, we implemented four out of five steps of Crisp DM process (Figure 1) - Data Understanding, Data preparation, Modeling and Evaluation - which will be extensively analyzed on the following section.

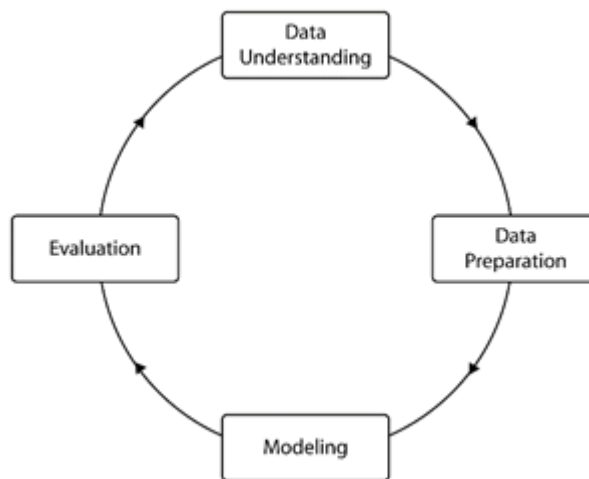


Figure 1: CRISP-DM methodology process

In order to initiate the aforementioned data analysis process, we collected network data using the NodeXL data importer, which allows access to social media and other forms of networks. In our analysis we imported data from "Twitter Search Network" because it allows the extraction of networks according to one or more hashtag(s). We used NodeXL client software to download Twitter data and selected only tweets that contain the hashtag "ekloges2015" which means elections of 2015 in Greek language, on a time frame of one week after Greek elections. The collected dataset consisted of the basic network, which includes published tweets and users' characteristics - such as user's "follows" or "mentions" or "replies" - along with the user's network of friends and the relationships between them. In order to achieve this research analysis' goal, additional data variables had to be used. Giving the limitations of NodeXL to get access on tweets' Retweets, Favorites and Comments, we directly requested data from Twitter's API. Using the Twitter's API console, we downloaded for the same date

range and the same hashtag the additional variables: Retweets, Favorites and Comments (Replies) for each observation.

4. Analysis and Results

4.1. Data Understanding

The data understanding phase starts with an initial data collection and proceeds with activities which aim to become familiar with the data, identify data quality problems, discover first insights and detect interesting subsets in the dataset. In the extracted network, vertices represent user entities while edges represent the interaction between them (Figure 2). Such interaction can be either a “follow” relationship between two users, a user “mention” in tweets, or a user “reply” to a tweet. Posted tweets are represented in the graph by a self-loop on the user who posted it. A “follows” edge means that one user follows another in the selected network. A “mentions” edge is created when one user mentions another user in a tweet (e.g. “being in the conference with @someone”). A “reply” edge is subtype of “follow” because it labels a relationship when one user refers to another at the beginning of the tweet (e.g. “@someone speaking right now”). Finally, a tweet is a simple post without a “reply” or “mention”. The entire network consists of 852 vertices and a total of 28.866 edges between them, in which 242 are tweets, 27.694 regard “follow” relationships and 930 are “mentions” to the tweets.

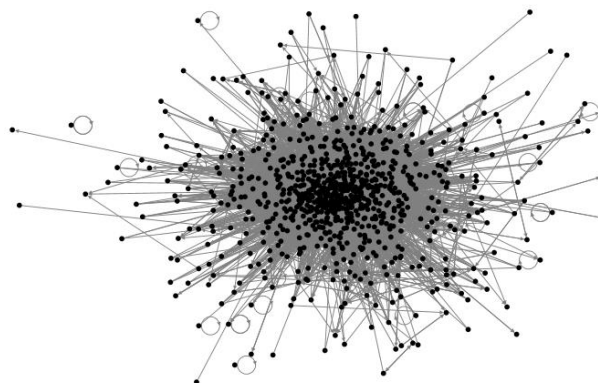


Figure 2: Extracted Twitter network

The structure of the initial data set we exploited, combines data from the two data sets and follows an entity - based approach. Thus, there are three distinct entities: “Users” who post content, “Tweets” which are the posted messages on Twitter and “Response” which is described by the actions users took on the posted tweets. Merging graph data with Twitter’s API export results, we composed a dataset with three different kinds of variables related to those entities. These data need to be exploited in order to create the useful for the analysis meta-data, as described in the following section. Our goal is to extract variables describing tweets characteristics - such as length, number of additional hashtags, number of URLs etc. -, attributes for user characteristics - such as number of followers, date of registration, total number of posts etc. - and data related to the response that each tweet got on Twitter - such as Retweets, Favorites and Mentions.

4.2. Data Preparation

The data preparation phase covers all activities to construct the final dataset from the initial raw data. As mentioned before, NodeXL provides the tweets - comments that are neither replies nor mentions- and some basic descriptive statistics such as the exact hashtags and URLs used. However, our research goal is to extract more useful meta-data like the tweets' length, number of URLs attached, number of hashtags used and number of mentions of other users. Thus, and within the entity - based approach, we isolated the tweets from NodeXL and the respective user's name that posted them.

Then we created the "Tweets Table" to store them and their most important attributes. Using appropriate SQL queries, we derived the following attributes: Tweet id (numeric), Content length (numeric), Number of hashtags of the tweet (including #ekloges2015), number of additional hashtags (other than #ekloges2015), number of attached URLs, number of mentions of other users. Merging the tweets' data with the Twitter's API export results, we expanded the dataset with three different kinds of variables related to those entities. Next, for each user that posted a tweet in this table, we isolated the useful data about them. These attributes are ID, Number of people that they Follow, Number of Followers, Number of Tweets posted, Number of Favorites.

According to the analytical approach that we chose to implement in this work, the dataset that will serve as input on the decided model should be properly be transformed. Our approach is to perform clustering analysis on tweets characteristics. To do so, we needed data for each tweet on whether it has a specific attribute or not. We coded this information into binary variables in where the value "1" represents the presence of this attribute in a tweet, and the value "0" represents its absence. However, we noticed that the tweets' variables were neither at the same scale nor on the same type in order to transform them into binary variables. To overcome this issue, we manipulated those attributes by scaling the numerical variables into classes. Then we either assigned the tweet on each class (value: 1) or not (value: 0). For example, the numerical attribute "Length" had a range from 35 characters (min) to 145 characters (max). We scaled this variable into 6 classes. Thus, in the final dataset a tweet that had 60 characters length was stored in the dataset in the following format:

| TweetID | length_class [35-53] | length_class [54-71] | length_class [72-89] | length_class [90-107] | length_class [108-125] | length_class[126-145] |
|---------|----------------------|----------------------|----------------------|-----------------------|------------------------|-----------------------|
| 100 | 0 | 1 | 0 | 0 | 0 | 0 |

4.3. Modeling

Having prepared the data and defined the fact table, we proceeded with the implementation of the clustering model. Our goal is to identify clusters of tweets' characteristics, meaning tweets that have common characteristics and thus can be grouped together. From our final dataset we used only the binary columns that represent tweets' attributes and consists of all types of data in the transformed classes which serve as the input of the model.

However, our research is based on a binary fact table so as to enhance the understanding of the results. Using binary data enables the researcher get results and understand each attribute value separately from the other values. In order to build the model, we used the x-means algorithm, which uses heuristically the k-means algorithm and can define the number of the deriving clusters without user's input. The algorithm uses a heuristic approach when creating the model and thus it is not a prerequisite to give the number of clusters as input. More specifically, the algorithm runs the model with different numbers of clusters on each iteration and ultimately selects the optimal number of clusters that maximizes the model's information gaining. At the end, it returns the possibilities of each attribute or attribute's value to appear in the respective cluster.

Implementing clustering algorithm on this dataset, we derived three clusters. The biggest cluster was "Cluster 1" accounting for the 68.95% of the total dataset, followed by "Cluster 2" which accounts for the 30.24% of tweets and "Cluster 3" which covers the remaining 0.81%. The strongest relation is shown between "Cluster 1" and "Cluster 3" in comparison with "Cluster 2". On the following section, we continue with the evaluation phase of the CRISP-DM methodology identifying more descriptive statistics about the clusters and concluding to some clusters' characterizations.

4.4. Evaluation

As the first step of the evaluation phase, we identified the main characteristics of each cluster. The clustering results provide us with the possibility of each attribute to appear in the tweets of the respective cluster.

Cluster 1: The main characteristic of the tweets in "Cluster 1" is that they all have at least one URL attached (100%). They also have 1 or 2 hashtags more that the examined one (#ekloges2015) (48%). Finally, the length of the tweet is from 126 up to 145 characters (46%), which means that they belong to "Big-sized" tweets. However, most of the characters in such tweets belong to the attached URL and not the message itself. In addition, only 1% of the tweets may "mention" another user. As a result, we label tweets of "Cluster 1" as "**Linked**".

Cluster 2: The first half of the tweets in this cluster has only the #ekloges2015 hashtag (51%), while the other half (47%) has 1 or 2 more hashtags. Furthermore, the tweets are medium (22% within the range 108 - 125 characters) to big sized (38% has from 126 up until 145 characters), which means that they have about 120 characters on average. Having no URLs attached and without mentioning other users, these tweets are characterized as "**Linguistic**".

Cluster 3: Tweets of this cluster have no other hashtags than the #ekloges2015 one (100%). In addition, more than half of them have a URL attached to the tweet (58%) and they are medium sized in length containing 72 to 89 characters (58%). Moreover, there is no mentioning of other users and thus these tweets are called "**Focused**". However, this cluster

contains only few tweets and thus it is statistically insignificant to continue with further analysis in it.

These characteristics are shown in the following figure (Figure 3):

| Variables | Population (All) | Cluster 1 | Cluster 2 | Cluster 3 |
|------------------|------------------|------------|------------|-----------|
| | | Multimedia | Linguistic | Focused |
| Length (108-125) | 42 | 32 % | 22 % | 0 % |
| Length(126-145) | 65 | 46 % | 38 % | 0 % |
| Length (35-53) | 5 | 1 % | 0 % | 21 % |
| Length(54-71) | 15 | 0 % | 15 % | 21 % |
| Length (72-89) | 19 | 5 % | 6 % | 58 % |
| Length (90-107) | 26 | 17 % | 18 % | 0 % |
| #Hashtags(0) | 91 | 42 % | 51 % | 100 % |
| #Hashtags(1-2) | 73 | 48 % | 47 % | 0 % |
| #Hashtags(3++) | 10 | 10 % | 3 % | 0 % |
| Mentions | 1 | 1 % | 0 % | 0 % |
| URLs in Tweet | 92 | 100 % | 0 % | 58 % |

Figure 3: Clusters' characteristics

After tweets clusters' characterization, we seek to understand what kind of response the tweets of each cluster received. As no tweet within the examined period got a reply, we will assess only the retweets and the favorites. In total, tweets in "Cluster 1" have collected 185 retweets and 149 "favorite" labels. On the other hand, tweets in "Cluster 2" have 97 retweets and 78 "favorite" labels respectively.

However, the clusters do not have the same size and hence cannot be objectively compared. As a result, we calculated the average number of retweets and favorites a tweet may get in each cluster. According to the results, a tweet that belongs to "Cluster 1" receives on average 1.16 retweets and 0.93 "favorite" labels, while a tweet in "Cluster 2" may get 1.33 retweets and 1.10 "favorite" labels. Interpreting the results, we can assume that "Linguistic" tweets are more likely to get a better response from their network, than "Linked" tweets.

The next step of our analysis focuses on the types of user profiles that post the tweets in each cluster. The user attributes that are available in our data set and represent the user's activity and behavior in Twitter are the number of people the user followed, the number of users that follow the user, the number of tweets the user posted in its entire history and the number of favorites it has. From these four attributes, we focus mainly on the total number of tweets the user tweeted, because it represents the intensity of the user's activity and the number of followers the user has, because it represents the maximum range of influence.

Users from "Cluster 1", i.e. users who posted tweets that belong to "Cluster 1", have a big number of followers and activity (i.e. tweets). However, users in "Cluster 2" exceed the first ones, by being very active and having the greatest number of posted tweets and followers among all three clusters. Moreover, we calculated the average of each metric and concluded with the following results. Users who post "Linguistic" tweets are the most active and followed, as they have the greatest number of tweets and followers, while users who post "Linked" tweets are the second most followed.

5. Discussion

This research is an exploratory work which aims to identify tweets' attributes related to its content, group them together according to their common characteristics and drill information about the impact of the tweet on the network. Using the tweets' characteristics, the tweet's impact on the network and the user's profile attributes, we aim to relate the clusters with the possible response they might get from the network. In addition, we aim to relate them with the types of users that post tweets of each cluster, in order to reach to some conclusions about the tweets' characteristics that may lead to bigger response rate for a typical Twitter user.

The results show that tweets can be grouped in two distinct clusters with unique characteristics that differentiate them from other groups. Some tweets focus on promoting a link to a video, a picture or another website ("Linked") and adding more hashtags to become visible, while others focus on communicating a message through a medium-length text in a tweet ("Linguistic") and mostly using one hashtag. Of course, there are also those tweets that point out their message via a small but targeted tweet ("Focused"), but they have not been examined further. Each type of tweet serves a different purpose, but also has a different impact on the network.

From the results, we derive that Twitter users tend to retweet more "Linked" tweets because they may contain interesting information and want to share it with their network. Such tweets reach many and sometimes different topics through the usage of more than one hashtags. On the other hand, they tend to favorite tweets with a long message, because they may be more meaningful for them. These results imply that if someone wants to be heard on the network, they have more possibilities to succeed if they accompany the message with a link to an interesting content, so as to indulge other users to share it with their network through retweets. This may be an effective way to make the message viral on the network. However, if someone wants to engage with their network, they should post long but meaningful content in order to entice others to favorite them. The results do not favor targeted but short tweets, which are neither favored nor retweeted. However, this cluster contains only a few tweets and thus no concrete conclusions can be derived.

The above-mentioned findings are enriched with some insights about the users who post different types of tweets. The most active users in terms of number of tweets post "Linguistic" tweets, while the most favored ones post "Focused" tweets and the most followed and the second most favored ones post "Linked" tweets. These results shed light from a different perspective as they clearly imply that the most followed users do not post "Linked" tweets, but tend to post medium-sized tweets in terms of content length and be favored more times. This could mean that they prefer to stay focused on one topic by using only one hashtag, promote one concrete message without the distraction of a link and build engagement with the network. We supposed this behavior is explained by the fact that known users with loyal followers do not need to use URLs, interesting articles and many hashtags to catch network's attention and create buzz.

Examining further these results, we looked in our data set to see, what type of tweets did the most famous users tweet. By famous or known users we refer to twitter accounts that represent influential individuals – like politicians, political analysts or journalists who are popular on society – and commercial accounts -like newspapers or blogs that became popular through digital social media. In order to achieve this, we identified the users that represent newspapers, famous blogs and news agencies in general. Such users tend to post more tweets that belong to Cluster 2, i.e. medium sized posts, with few hashtags and no URLs. However, there are also many online news agencies that post tweets mainly with URLs attempting to advertise their blog or website, and thus belong to Cluster 1. As a result, we concluded that there are no clear trends from the famous users.

6. Limitations - Future research

NodeXL Twitter Search network operator communicates with Twitter service in order to import data. Our query targeted 1000 tweets which were easy to be found because of this hashtag's popularity (our hashtag was #ekloges2015 and the data collection happened right after the Election Day). However, we imported not only the tweets, but also the network for each user who posted the tweet. Thus, NodeXL downloaded only 250 tweets, a relatively small dataset, setting the first limitation of our research. A small dataset narrows the scope of our research and can be biased as it is not representative of the total population. Future research will examine a different dataset, with larger number of tweets which will be extracted using the Twitter API. Using the latter method we could target specific fields (e.g. tweet, tweet_id, user, number of followers etc.) which are needed for our research and belong to a predetermined time frame. Extracting less but more useful information would lead to collecting a much bigger dataset in less time.

Our research focuses on tweets about a specific occasion, the Greek national elections of 2015. Hence the results are applied only to this topic and we cannot assure external validity to any other type of event, occasion or topic. In addition, tweets were extracted from a timeframe of seven days after the elections had finished and thus do not represent users' behavior throughout the elections campaigns when tweets may get different response - more retweets or replies. Additional work could compare analysis results from both before elections date and after. More specifically, our next step is to select a specific elections' campaign beforehand, so as to collect tweets a few weeks before the elections, throughout the elections' day and a few weeks afterwards. In this way, we will cover the whole elections' period and be able to make comparisons and more in depth analyses.

Moreover, our current study uses exploratory method on social network analysis through a heuristic approach, using k-means, a heuristic algorithm, which is not without its limitations. First of all, it cannot work with categorical data, but only numerical values (Huang, Zhexue, 1998). In addition, another limitation of this clustering algorithm is that it cannot handle empty clusters and outliers, while the researcher has to reduce SSE with post-processing tasks (Singh et al, 2011). Future studies could use another clustering algorithm to compare the results or switch to a different analysis approach. We are currently extending this work by

switching on explanatory analysis aiming to identify the most important tweets' attributes that lead to intense response from the network and how we can predict response outcome based on tweets' and users' characteristics.

Another limitation derives from the scope of the current research, which focuses on the contribution of tweet's attributes on the impact a tweet has on its network. However, social sciences show that user can affect tweet's impact. Thus, in future studies we will examine if the popularity of the user may influence the response a tweet has on its network. In order to define objectively the term 'popularity', we will extract tweets from the top tweeters who used the relevant hashtag. Comparing their impact to the less 'popular' users' impact, we could enhance our analysis and answer our hypothesis.

Moreover, this study examines the impact of shared content (tweet) in Twitter, which is only one from many ways to interact with a network in available social media platforms. Future research could compare which attributes relate to the influence of a message in different social platforms. In specific, we could extract posts (e.g. tweets from Twitter, posts from Facebook etc.) about the same topic and compare their attributes among different platforms. Finally, in addition to the above-mentioned findings, it would be interesting to identify whether the big network of followers leads to more retweets of the message or the many retweets lead to a bigger network of followers.

7. References

- Boyd, danah, Scott Golder, and Gilad Lotan. 2010. "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter." HICSS - 43. IEEE: Kauai, HI, January 6
- Bruns, Axel, and Stefan Stieglitz. "Quantitative approaches to comparing communication patterns on Twitter." *Journal of Technology in Human Services* 30.3-4 (2012): 160-185.
- Bruns, Axel, and Jean Burgess. "# ausvotes: How Twitter covered the 2010 Australian federal election." (2011): 37.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., & Shearer, C. "1.0 Step-by-step data mining guide, CRISP-DM Consortium, 2000."
- Cheung, Kwok-Wai, et al. "Mining customer product ratings for personalized marketing." *Decision Support Systems* 35.2 (2003): 231-243.
- Giudici, Paolo, and Gianluca Passerone. "Data mining of association structures to model consumer behaviour." *Computational Statistics & Data Analysis* 38.4 (2002): 533-541.
- Go, Alec, Richa Bhayani, and Lei Huang. "Twitter sentiment classification using distant supervision." CS224N Project Report, Stanford 1 (2009): 12.
- Griva, Anastasia, Bardaki, Cleopatra, Panagiotis, Sarantopoulos, Papakiriakopoulos, Dimitris,, "A data mining-based framework to identify shopping missions" (2014). MCIS 2014 Proceedings.
- Harrison, Tina S. "Mapping customer segments for personal financial services." *International Journal of Bank Marketing* 12.8 (1994): 17-25.
- Huang, Zhexue. "Extensions to the k-means algorithm for clustering large data sets with categorical values." *Data mining and knowledge discovery* 2.3 (1998): 283-304.
- Kwan, Irene SY, Joseph Fong, and H. K. Wong. "An e-customer behavior model with online analytical mining for internet marketing planning." *Decision Support Systems* 41.1 (2005): 189-204.

- Pennacchiotti, Marco, and Ana-Maria Popescu. "A Machine Learning Approach to Twitter User Classification." *ICWSM 11* (2011): 281-288.
- Shi, Zhan, Huaxia Rui, and Andrew B. Whinston. "Content sharing in a social broadcasting environment: evidence from twitter." Available at SSRN 2341243 (2013).
- Singh, Kehar, Dimple Malik, and Naveen Sharma. "Evolving limitations in K-means algorithm in data mining and their removal." *International Journal of Computational Engineering & Management* 12 (2011): 105-109.
- Stieglitz, Stefan, and Linh Dang-Xuan. "Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior." *Journal of Management Information Systems* 29.4 (2013): 217-248.
- Tumasjan, Andranik, et al. "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment." *ICWSM 10* (2010): 178-185.

EVALUATING PRIVACY PRACTICES IN WEB 2.0 SERVICES

Completed Research

Vemou Konstantina. University of the Aegean, Department of Information and Communication Systems Engineering, Samos, Greece, kvemou@aegean.gr

Karyda Maria. University of the Aegean, Department of Information and Communication Systems Engineering, Samos, Greece, mka@aegean.gr

Abstract

This paper discusses the effectiveness of privacy practices and tools employed by Web 2.0 service providers to facilitate users protect their privacy and respond to public pressure. By experimenting on three recently introduced tools, which claim to offer users access and choice on the data stored about them, we analyse their privacy preserving features. Research results indicate their limited effectiveness with regard to user privacy. We discuss discrepancy between stated goals of these privacy enhancing tools and actual goals these tools accomplish.

Keywords: Privacy practices, privacy communication game, effectiveness, fair information principles.

Introduction

In the Web 2.0 era users are the centre of internet services, providing and sharing vast amounts of information. Whether publishing their status on Social Networking Services (SNS), or browsing websites, communicating or conducting commercial transactions, users entrust service providers with large amounts of personal data. This data can be used to assist users in their internet experience, by providing personalized services (Berger 2011) or optimizing the results of internet searches (Cooper 2008). However, gathered user data also expose users to a set of privacy threats, such as price discrimination (Gross and Acquisti 2005), out of context use of information (Nissenbaum 2009) and revelation to unintended audiences (Van den Berg and Leenes 2010). Users' concerns over privacy are rising (Acquisti and Gross 2006; Karyda and Kokolakis 2008; Boyd and Hargittai 2010; Pavlou 2011); however they are continuing to share their personal data, mainly on SNS. It has also been pointed that users are not aware of several privacy invading practices, such as third-parties tracking their online activity (ENISA 2012). Lately several incidents of privacy violation have reached the spotlight, alarming online users. For instance, the introduction of the Newsfeed feature in Facebook, which delivered profile updates to other users in headline news with limited audience settings (Hoadley et al. 2010) and the Google's circumvention of barriers in Apple's Safari browser to track users' online activity have reached the mass media, creating a public outcry from privacy advocates (Gorman 2012). Another example in the long range of published privacy violation incidents was the leak of Snapchat user photos (Snapchat is a mobile phone app to send self-deleting photos), via hacking in "Snapsaved.com", a service offering use of Snapchat on a desktop computer. As reported, Snapsaved.com had failed to adequately inform users on the saving practice of their pictures (The Guardian 2014).

Digital media providers are profit organizations with an interest in maintaining highly active users (Easley and Kleinberg 2010; Krasnova et al. 2009). They manage users' personal data to achieve business goals, including enhanced products and services and more effective and efficient operation. However, with privacy advocates reproving their privacy practices after each violating incident on the

news, companies are forced to seek for solutions to appease awakened privacy concerns. Popular social media platforms have lately introduced a variety of settings, including access controls of posted information and activity logs and applied changes on their privacy policies. According to Pollach (2005) such initiatives mainly occur as a response to international law compliance needs and public pressure from privacy advocates and the media.

The effectiveness and motives of these practices have been the subject of debate. For instance, in May 2015, 80 academics sent an open letter to Google requesting more transparency in the process of citizen requests for “their right to be forgotten”, which is requests of delisting links to personal information that is inaccurate, inadequate, irrelevant, or excessive for the purposes of data processing (The Guardian 2015). Hong (2014) argued on the privacy placebos; tools or settings that allay privacy concerns, but gain limited user acceptance or offer little actual value on protecting user privacy. Bonneau and Preibusch (2010) refer to this as the “privacy communication game”, where service providers are struggling to provide privacy practices that will appease privacy concerns for privacy-aware users, without adequately addressing them, and without alarming the non-privacy-aware, thus creating a back door to continue taking advantage of user personal data.

In this paper we further explore this debate, and investigate whether recently adopted privacy practices are targeted to protecting users’ privacy or if they serve different purposes, as part of a communication game, for instance. Using a set of criteria derived from the Fair Information Principles (Schwaig et al. 2006), we study the effectiveness in protecting user privacy of three tools recently introduced by popular providers: *Google search history*, *Google takeout* and *Facebook download your data archive*. The remaining paper is structured as follows: in chapter 2 we present the debate on privacy practices effectiveness. In chapter 3 and 4 respectively, we present the criteria as well as the analysis we performed on the above mentioned tools, followed by a discussion on their effectiveness in chapter 5 and conclusions in chapter 6.

2 Privacy Practices in Web 2.0 and their Effectiveness

Responding to negative publicity and user concerns over personal data practices, Web 2.0 service providers apply a wide range of privacy practices, including publishing privacy policy documents and introducing privacy settings. Pollach (2005) argues that the introduction of privacy practices is driven mainly by the need to comply with privacy laws and the need to gain users’ trust, by recognizing their concerns over privacy. Starting from the early 1990s when e-commerce began to flourish, privacy policies have been used as a means for companies to inform online users of their privacy practices. The quality of privacy policies in terms of guaranteed privacy levels (adhering to fair information principles) has been linked to users’ trust and engagement with the web service providers (Lauer et al. 2003; Meinert et al. 2006). However, the effectiveness of privacy policies in protecting users’ privacy has been questioned by many researches. There may be many reasons explaining why users ignore the website’s privacy policy, including its length, obscure language used (Lauer et al. 2003; Pollach 2005), or the difficulty to access it in the website (Bonneau and Preibusch 2010). Other studies found that providers tend to describe certain privacy practices in their policies, such as the collection of login information, but omit documenting other practices, such as informing users of third-parties their data will be shared with (Pollach 2005; Stuart 2007).

Several platforms employed the Platform for Privacy Preferences (P3P) (Cranor 2003), which, however, did not gain widespread acceptance and has been gradually abandoned. The idea behind P3P was that websites would express their privacy practices in a standard format and this would automatically be compared to users’ privacy preferences, to assist them in their decision to browse or not to browse the website. P3P’s fall backs include partial address of the fair information principles as it did not provide any mechanism to assure companies process personal data according to stated privacy policies (enforcement) (Jaatun et al. 2011) and the ability of companies to circumvent user’s preferences by providing erroneous or conflicting P3P compact policies (Leon et al. 2010).

Privacy seals, such as TRUSTe, BBBOnline and WebTrust, are third-party programs assuring that an organization's handling of personal data is aligned to privacy practices declared in its privacy policy. In general they are a forcing task for organizations to provide a privacy policy, as well as to align it to fair information principles and are used to convince users that their privacy is not violated by using their online services. They have also received criticism, on the grounds that privacy seals institutions do not oblige organizations to explicitly mention all types of information collected and how they will be used. Also, as they are targeted at ensuring compliance with the company's privacy policy, there are workarounds to using personal data, by declaring them as assets in the privacy policy (Moore and Dhillon 2003). What is more, in the event of privacy breaches, the assuring institutions can only revoke the seals and cannot help users in regaining their lost privacy (Shapiro and Baker 2001; Hui et al. 2007).

Opt-in strategies requiring that profilers obtain consent prior to collecting or sharing user data (Berger 2011) have also been debated as they are applicable to only binary policies (accept or not accept to provide data), while the privacy needs of users differ according to the purpose of personal data use (Spiekermann and Novotny 2015). It was also argued that companies can coerce users in accepting their practices by refusing to provide their services in any other case (Berger 2011; Cooper 2008). At the same time, as privacy threats on certain web 2.0 services, such as SNS, are not only based on the platform's tracking of user behaviour but also on the difficulty to protect semi-public information published by the users themselves, opt-in or opt-out practices are only partially covering users' privacy protection needs.

Finally, privacy settings, such as access controls, introduced in several SNS platforms, are usually judged as insufficient or unusable and were found to be neglected by users (Boyd and Hargittai 2010). This is alarming, as default privacy settings were found to be non-conforming to privacy-by-design principles (Vemou and Karyda 2014) and tend to become progressively more liberal over time (Michota and Katsikas 2014).

3 Theoretical Background

3.1 Research approach

This paper investigates if privacy practices adopted by Web 2.0 service providers are targeted to protecting users' privacy or whether they pursue different goals. To gain further insight into this issue we performed an experimental analysis, using several accounts on three tools: *Google Takeout*, *Google Search History* and *Facebook download your data archive*.

These tools were selected based on the popularity of the platforms (Statistics by Alexa 2015) and on their functionality: they are mainly targeted at providing users with the ability to access/control information these two companies gather about them. Thus they are focused on providing users with control over their personal information, which has been a major privacy requirement by users. These tools have been recently launched and have not been evaluated in relevant research.

3.2 Conceptual framework

The evaluation of privacy enhancing tools was based on criteria (described in Table 1) derived from the set of fair information practices (FIPs), introduced during the 1990s by US Federal Trade Commission (Schwaig et al. 2006), as well as from relevant literature, extending them (Anton et al. 2002; Spiekermann and Cranor 2009).

| Fair information principles | Evaluation criteria employed in this paper |
|---|---|
| <p><u>Notice/awareness</u> Companies need to inform users if and which personal information is being collected and how such information will be used.</p> | <ol style="list-style-type: none"> 1. Does the tool provide users with a full list of collected and shared personal information? 2. Does the tool provide users with a detailed list of third-parties information will be shared with? 3. In which form is users' data shared with third parties (anonymized, aggregated etc.)? 4. Does the tool inform users of privacy threats stemming from collected and shared personal data? |
| <p><u>Choice/consent</u> When it comes to use of gathered personal information, users should be provided with the choice of use purpose</p> | <ol style="list-style-type: none"> 5. Does the tool enable users to declare preferences on how their data will be used? 6. Does the tool enable users to declare preferences on whether this information will be shared with third parties (unless required by law)? 7. Does the tool inform users of personal data processing for purposes they have not consented? 8. Does the tool inform users of personal data processing by third-parties, in cases of their choice to prohibit it, despite prior data sharing? |
| <p><u>Access/participation</u> Users have the right to access information about them and to correct errors.</p> | <ol style="list-style-type: none"> 9. Does the tool allow modifications/corrections on collected personal data? 10. Does the tool enable users to request complete deletion of their data? 11. Does the tool enable users to access and correct data processed by third-parties? |
| <p><u>Integrity/security</u> Companies should take preventive measures for unauthorized access during transmission and storage.</p> | <ol style="list-style-type: none"> 12. Does the tool prevent other users from gaining access to information stored about a certain user? |
| <p><u>Enforcement/redress:</u> Companies should be able to demonstrate operation as dictated by stated privacy policies.</p> | <ol style="list-style-type: none"> 13. Is it visible to users (via the tool) that data processing is performed according to their declared choice? 14. Is it visible to users (via the tool) that stored information is actually corrected or deleted following their requests? |

Table 1. Evaluation criteria used

4 Evaluating Privacy Tools that Provide Access to Stored Information

In this section we evaluate three tools that claim to provide users with access, transparency and choice on personal data, developed by the two most popular Web 2.0 service providers, Google and Facebook.

4.1 Google Search History

“Google Search History” is a service provided to Google users, as part of Google’s Dashboard, presenting their search history in the Google Search engine. Users are presented with their searches performed by date and time, and the pages visited from the respective search results. Also, graphical rep-

resentations of search activity statistics are provided, such as daily or weekly activity rates, and other statistics, such as the most frequent searches of the user.

According to Google, this tool offers transparency and provides users with control over the data associated with their accounts¹. This tool is also included in the privacy policy (in the section “transparency and choice”), as part of a set of privacy preserving services. Use of the tool requires logging into any Google service, to ensure security and privacy of presented information.

In general, this tool raises user awareness on collected data from search history. However, presented information is limited to users’ searches, although information on their other activities is gathered and, if combined with search history, can lead to their online profiling. What is more important is that such information gathering is not transparent to the users and does not always require their login to Google services (e.g. tracking on other webpages).

Other information, such as the location of the users or the time spent browsing result pages, could be of significance if correlated to a search; however users do not have access to this information. Also, users could form a false assumption that information on their location is not gathered and correlated to their searches, in case they have chosen not to share their location. In this case they would be presented with an empty locations history report (in the account history settings), albeit according to its privacy policy, Google maintains IP records from their logins. While this could automatically reveal users’ location (in area level) and could be correlated to their searches, the search history tool fails to provide users with access to this inferred information.

This tool does not inform users on how data on their search activity is collected or shared with third-parties, offering no traceability of such data and no choice on how it will be used. For instance, no privacy signalling technique is offered to declare accepted purpose of collected data use, and there is no opportunity to correct or delete data stored in third-party platforms. As for the search data stored by Google platform, users are given the opportunity to delete past searches from their search history or to disable collection of search history elements, but there is no transparency on whether collected data are actually deleted as result of the users’ request or they just do not appear in the history tool interface. Thus “*Google Search History*” tool addresses notice and access principles up to an extent, but fails to provide users with choice on how their data is handled.

4.2 Google Takeout

“*Google Takeout*” is a tool provided to Google users, in order to download data they have imported or created in several Google websites/services. According to Google, it provides users access to their information, as stored by the company and can also serve for backup of user data or for importing it to other services². It can also serve as an awareness raising tool, as users’ access to their personal data is of critical importance for privacy protection. Other platforms, such as Facebook, offer similar services as privacy preserving practices.

Users need to be logged in their accounts, in order to access “*Google Takeout*” from the account settings and when they try to access/download the archive they are prompted to re-enter their account credentials, although they may follow a link sent to their email account.

“*Google Takeout*” offers users the opportunity to download a copy of their imported data from several Google services such as Gmail, Contacts, Google+ and Google Latitude. In the created archives, users are presented with the information they have inserted or directly created as a result of using these services. For instance, users are presented with a history of YouTube videos they have watched. However, provided information does not include other information the company has stored on users, such as

¹ <http://www.google.gr/intl/en/policies/privacy/>

² <https://support.google.com/accounts/answer/3024190?hl=en>

log data. An exception to this is the location history of the users, which presents them with a history of the places they have been (in coordinates), along with inferred activities, e.g. being in a vehicle or walking. Even in this case, other inferred and stored information, such as identification of work or home places, is not provided.

Despite offering a presentation of users' stored data, "Google Takeout" fails to communicate which entities have access to this data, other than the Google Platform and the users themselves. Users are not presented with information derived from their data, as a result of processing and combining different services use data, while relation of user data to privacy threats, such as price discrimination or publishing to unwished audiences is omitted.

"Google Takeout" does not provide users with an interface to report their preferences on how their data is data used. Also, although "Google Takeout" serves as single point users can access stored information from different Google services, it is ineffective in directly requesting correction or deletion of some data. Instead, users need to access the different services websites to modify or delete them. Concluding, "Google Takeout" enhances users' awareness.

4.3 Facebook download your data archive

"Facebook download your data archive" is a tool provided by Facebook for users to get a copy of their data as stored by the platform. As stated in Facebook's privacy policy³, this is a feature to enable users' access to stored information about them. The downloaded archive contains a list of user created and posted data, such as profile information, photos and friends list, but also other categories of data stored without users' notice, such as log data and advertisement activity (e.g. which advertisements the user has clicked).

The tool makes users aware that data not directly created by them are also stored by the company. For instance, the archive of user data contains friend requests, either generated by the users themselves or by other users, regardless of acceptance. Also login data (device and IP addresses) and information on advertisements users have clicked are presented. In the same way, users are presented with some basic information on data inferred by their activity and posts, such as advertisement categories they are targeted with. However, there is no reference to other types of inferred data from the users' activity, which audiences got access to users' activity (e.g. wall posts), which data were used in advertisements displayed to the users' network and which were shared to third-parties. Furthermore, the network created by friends of friends is not graphically or in any other way presented, although it is stored and can be exploited by the company.

Although the feature is mentioned in the part of privacy policy concerning users' access and correction of data stored about them, the only automated option to correction/deletion from the data archive is to delete user account. Even in this case, as stated, some parts of information, such as logins will be maintained by the platform, for at least a year. In case of content deletion from the account, e.g. removal of a photograph or video, this will be removed from the downloadable archive, but users have no direct means to confirm that the item itself, as well as metadata gathered by the platform are deleted and also have no means to enforce such deletion on third-parties that have accessed it. "Facebook download your data archive" covers a range of notice requirements, by providing users with secondary information stored about them, but fails to provide effective choice and access mechanisms.

³ <https://www.facebook.com/privacy/explanation>

5 Discussion

The tools analysed in this study aim to provide users with access, transparency and choice with regard to their personal information gathered by Web 2.0 providers. Research results, based on Fair Information Principles showed that these tools succeed in informing users of stored data about them, which is a step towards notice and raising awareness. However, we found that not all information stored about users is presented, and users are not presented with the outcome of processing such data. For instance, in Google's tools, log data were omitted, although the knowledge of IP address of users' connection enables the company to presume their location. Also, inferred information on the users' networks, such as proximity of friends based on the frequency of their communications or the combination of their locations was not presented in Facebook's tool. Information on third-parties that gained access to users data were not presented via the studied tools, and there was no report on how stored or published information could lead to privacy threats for the users.

Providing users with choice over their personal information was also among the stated goals of these tools; however our analysis revealed that the tools did not offer any method for users to present their preferences on how their data should be treated. For instance, users may wish their interests or hobbies information to be used for improved search results or for friend recommendations, and not for marketing purposes, but there would be no choice to express it via the analysed tools.

Moreover, users were not presented with the choice to request correction for data stored of presumed about them. For instance, while Facebook's archive informed users of the advertisement categories they were enlisted, based on their activity, they were not provided with the ability to opt-out of a certain category of advertisements. In some cases users were provided with the capability to delete some parts of stored information (e.g. from google history), but were not in the position to control whether this information or related information (e.g. from cookies) were actually deleted from the company. Apart from this, in some cases, such as in Facebook archive, users were only offered with the choice of complete deletion of the account.

Typically, platforms provide several privacy enhancing tools to address complementary privacy requirements of users. Based on our analysis, even if more than one tool were used, certain aspects of user privacy would still be unaddressed. For instance, the application of *Google search history* along with *Google Takeout* would not cover requirements related to choice of data use purpose. In the same context, Google's Dashboard provides no options to declare preferences on how their data can be used. Finally, a combination of the above tools could not provide a more detailed list of third-parties gaining access to user data, thus notice/awareness principle would still be partially addressed.

Another issue derived from our analysis is the presentation of such tools to users. In Facebook, the tool to download user archive was placed at the end of the account settings list, in smaller font than the other settings. It seems that it was intended to separate this tool from other settings, alas highlight its existence. This could be viewed as a usability issue, posing questions on the platform's motives of introducing the tool. Also, our experience with these tools shows that usability in terms of presentation of data could be significantly improved. An example of such improvement is to present formed networks of friends or the content of advertisements users have clicked. Also, *Google Takeout* used open archive templates (JSON etc.) to help users import their data to other platforms, but it would be preferable to also provide a way of accessing stored data in webpages, for users with less technical skills.

In conclusion, our analysis revealed a discrepancy between stated privacy goals and achieved goals of the mentioned tools. They were found to cover only partial requirements on notice and awareness, while superficially treating other privacy preserving principles, such as access/participation. The last was mainly stated as their goal. Also, the analysed tools did not address choice/consent and enforcement principles. Usability preservations led to questions on actual reasons the tools were introduced for, taking into account that the majority of users are not IT specialists and non-usable or difficult-to-locate tools could lead to their avoidance. In fact, companies do not seem to be paying much

attention on usability, but this could be part of their strategy playing the “privacy communication game”.

6 Conclusions and Further Research

This paper explores a “privacy communication game”: companies adopting privacy practices or introducing tools to help users in their efforts to protect their privacy and appease public outcries, while in fact their effectiveness or usability is limited. By evaluating three recently introduced tools that promise users access and choice on data stored about them, and analysing their features based on the Fair Information Principles, we reveal their limited effectiveness with regard to privacy protection, even in cases of their combined use. Our analysis shows that there may be a discrepancy between the stated goal, which is to protect users and offer them rights on their data, and actual goals of the companies, such as appeasing public pressure and marketing privacy to temper users’ privacy concerns.

Web 2.0 service providers employ a business model that partly depends on the exploitation of user data to derive its profits. Privacy practices and tools, while presented as solutions to users’ privacy concerns, are largely used as a reactive “damage control” instrument, after privacy breaches or invading practices reach the spotlight and receive criticism. We also found the usability of the privacy tools we studied rather low, which can be explained, as they are meant to offer some control and appease concerns of privacy-aware users but in such a way that non-privacy-aware users do not understand privacy threats they are exposed to.

As a result, available privacy tools are in some cases extending the privacy illusion users are under, despite promising control over their personal data. In an attempt to raise users’ awareness and assist them, a set of criteria to evaluate if tools meet their privacy requirements, regardless of what is stated or communicated by the platforms should be offered. Also, as awareness of PETs existence is found to be of importance for users to engage with them, and as it is a common practice for offered tools to be “hidden” in account settings, actions to raise awareness of the tools existence are also necessary.

References

- Alexa Internet, Inc. (2015). “The top 500 sites on the web”, URL: <http://www.alexa.com/topsites/> (visited on 09/06/2015).
- Antón, A., Earp, J. B., and Reese, A. (2002). “Analyzing website privacy requirements using a privacy goal taxonomy”. In: Proceedings of IEEE Joint International Conference on Requirements Engineering. Ed. by Danielle C. Martin, Essen: Germany, p. 23-31.
- Berger, Dustin D. (2011). “Balancing consumer privacy with behavioral targeting”. *Santa Clara Computer and High Technology Law Journal* 27 (2011): 3.
- Bonneau, J. and Preibusch, S. (2010). “The privacy jungle: On the market for data protection in social networks”. *Economics of information security and privacy*. Springer US, p. 121-167.
- Boyd D. and Hargittai E. (2010). “Facebook privacy settings: Who cares?”. *First Monday*, 15(8).
- Cooper, A.(2008). “A survey of query log privacy-enhancing techniques from a policy perspective.”, *ACM Transactions on the Web (TWEB)* 2.4 (2008): 19.
- Cranor, L. (2003). “P3P: Making privacy policies more useful”, *IEEE Security & Privacy*, 1(6), p.50 - 55.
- Easley, D. and Kleinberg, J. (2010). “*Networks, Crowds, and Markets: Reasoning About a Highly Connected World*”, Cambridge University Press, Cambridge
- ENISA (2012). “*Privacy considerations of online behavioural tracking*”. Report 2012
- Gorman, G.E. (2012). “From deception to deceit: Google and privacy? Don’t make me laugh”, *Online Information Review*, 36(3).

- Gross, R., and Acquisti, A. (2005). "Information revelation and privacy in online social networks" In: Proceedings of the 2005 ACM workshop on Privacy in the electronic society. Alexandria: USA, p. 71-80.
- Hoadley, C., M., Xu, H., Lee, J. J., Rosson, M. B. (2010). "Privacy as information access and illusory control: The case of the Facebook News Feed privacy outcry". *Electronic commerce research and applications*, 9(1), p. 50-60.
- Hong, J. (2014). "Research Issues for Privacy in a Ubiquitously Connected World", Technical Report, NITRD National Research Strategy on Privacy, 2014.
- Hui, K.L., Teo, H.K., Lee, S.Y.T. (2007). "The value of privacy assurance: an exploratory field experiment." *MIS Quarterly* (2007), p.19-33.
- Jaatun, M. G., Tøndel, I. A., Bernsmed, K., Nyre, Å. A. (2011). "Privacy enhancing technologies for information control". *Privacy Protection Measures and Technologies in Business Organizations*, p. 1-31.
- Karyda, M., Kokolakis, S. (2008). "Privacy Perceptions Among Members of Online Communities, In: *Digital Privacy: Theory, Technologies and Practices*. Ed. by A. Acquisti, S. De Capitani di Vimercati, S. Gritzalis, C. Lambrinoudakis. Auerbach Publications (Taylor and Francis Group)
- Krasnova, H., Hildebrand, T., Guenther, O. (2009). "Investigating the value of privacy in online social networks: conjoint analysis." *ICIS 2009 Proceedings* (2009): 173.
- Lauer, T. W., Deng, X. (2007). "Building online trust through privacy practices". *International Journal of Information Security* 6.5 (2007), p. 323-331.
- Leon, P. G., Cranor, L. F., McDonald, A. M., McGuire, R. (2010). "Token attempt: the misrepresentation of website privacy policies through the misuse of p3p compact policy tokens". In: Proceedings of the 9th annual ACM workshop on Privacy in the electronic society, Chicago: USA, p. 93-104.
- Michota, A. and Katsikas, S. (2014). "The evolution of privacy-by-default in Social Networks". In: *Proceedings of the 18th Panhellenic Conference on Informatics*. Athens: Greece.
- Moore, T. T., Dhillon, G. (2003). "Do privacy seals in e-commerce really work?" *Communications of the ACM - Mobile computing opportunities and challenges*, 46(12), p. 265 – 271.
- Meinert, D.B., Peterson, D.K., Criswell, J.R., Crossland, M.D. (2006) "Privacy policy statements and consumer willingness to provide personal information". *Journal of Electronic Commerce in Organizations*, 4(1), p. 11–17.
- Nissenbaum, H. (2009). "Privacy in context: Technology, policy, and the integrity of social life". Stanford University Press.
- Pavlou, P. (2011). "State of the information privacy literature: where are we now and where should we go?" *MIS quarterly*, 35(4), p. 977-988.
- Pollach, I. (2005). "A typology of communicative strategies in online privacy policies: Ethics, power and informed consent". *Journal of Business Ethics* 62(3), p. 221-235.
- Schwaig K.S., Kane G.C., Storey V.C.. (2006). "Compliance to the fair information practices: How are the Fortune 500 handling online privacy disclosures?", *Information and Management*, 43 (7) , p. 805-820.
- Shapiro, B. and Baker, C.R. (2001). "Information technology and the social construction of information privacy". *Journal of Accounting and Public Policy*, 20(4), p. 295–322.
- Spiekermann, S., Cranor, L. F. (2009). "Engineering privacy". *IEEE Transactions on Software Engineering*, 35(1), p. 67-82.
- Spiekermann, S., and Novotny, A. (2015). "A vision for global privacy bridges: technical and legal measures for international data markets". *Computer Law & Security Review* 31(2), p. 181-200.
- Stuart, A.H. (2007) "Online Privacy Policies: Contracting Away Control Over Personal Information?." *Penn State Law Review* 111 (3) (2007).
- The Guardian. (2014). "Snapchat videos and pictures stored on a third-party website posted online", URL: <http://www.theguardian.com/technology/2014/oct/12/teenagers-snapchat-images-leaked-internet> / (visited on 09/06/2015).

- The Guardian. (2015). “Dear Google: open letter from 80 academics on 'right to be forgotten’”, URL: <http://www.theguardian.com/technology/2015/may/14/dear-google-open-letter-from-80-academics-on-right-to-be-forgotten> /(visited on 09/06/2015).
- Van den Berg, B. and Leenes, R. (2010). “Audience segregation in social network sites”. In: Proceedings of the IEEE Second international conference on Social Computing (SocialCom).
- Vemou K. and Karyda, M. (2014). “Guidelines and tools for incorporating privacy in Social Networking Platforms”. *IADIS International Journal on WWW/Internet* 12(2), p. 16-33.

PRIVACY AND THE DIGITAL GENERATION GAP: MYTH AND REALITY

Complete Research

Ruth Halperin, College of Management Academic Studies, Israel, ruth.halperin@gmail.com

Yuval Dror, College of Management Academic Studies, Israel, yuvald@colman.ac.il

ABSTRACT

Over the past decade the demise of privacy has been repeatedly pronounced by renowned technology executives such as Mark Zuckerberg declaring privacy as passé and anachronistic- “so 20th century” - the concern of old people. However, there has been relatively little research into privacy perception and behaviour among different generations that may relate to how people navigate their private lives in online settings. Furthermore, recent research reveals the ways in which privacy concerns of young internet users are enacted, thus challenging overgeneralized claims of a clear-cut generation gap associated with online privacy. As information privacy problems are becoming thornier, unfounded statement voiced by stakeholders with vested interests should be put to one side. Instead, systematic research is needed to understand how privacy is perceived and managed by people of different age groups, and what measures can and should be taken to address current and future concerns of internet users across generations. We address these questions and account for the results using a representative sample from Israel.

Key words: privacy, self-disclosure, generations, digital natives

1 Introduction

Information privacy is considered one of the most important ethical, legal, social, and political issue of the information age (Hong & Thong, 2013; Dinev, 2014; Acquisti et al, 2015). During the past decade, the problems involving information privacy—the ascendance of Big Data and fusion centers, the tsunami of data security breaches, the dramatic rise of Web 2.0, the growth of behavioral marketing, and the proliferation of tracking technologies — have become thornier (Solove, 2013: 1). Personalized web services and business intelligence software require the collection and mining of unprecedented amounts of personally identifying information. As consumers become content providers on web blogs and social networking websites (SNS), their personal information becomes ever more vulnerable (Hong & Thong, 2013). Researchers are advocating a re-examination of privacy concerns and consequences to reflect the contemporary nature of this dynamic construct (Anton et al, 2010). Under these conditions, understanding individuals' privacy concerns and behaviour is fundamental to the success of emerging digital technologies (Pavlou, 2011). In this paper, we investigate the differences in privacy concerns and behaviour between young and old generations. While many popular claims are in circulation (Nussbaum, 2007), findings on this question are limited at best (Steijn & Vedder, 2015). It is therefore important to establish based on rigorous, representative data, the effects of age on privacy related perceptions and practices and to increase understanding of actual generational divides to inform legislators and policy makers, as well as internet entrepreneurs (Steijn & Vedder, 2015; Wang et al, 2013).

2 Online Privacy and the Digital Generation Gap

The publication of Mark Prensky's article entitled 'digital natives, digital immigrants' (2001) gave rise to a fierce debate of the relationship between IT and age. Prensky argued that digital technology has brought about a radical generation break between young people, so-called 'digital natives', and the older generation of 'immigrants'. In line with this trend, and albeit the critical wave (cf. Herring, 2008; White and Le Cornu, 2011) a generational rhetoric entered the discourse about online privacy (Barnes, 2006). As social media proliferate, the prevailing notion persists: young people don't care about privacy while older generations still do (Tufeky, 2012; Boyd, 2014).

Yet, research to corroborate the claims is scarce. Few empirical studies can confirm that young people actually express fewer privacy concerns than adults (Steijn & Vedder, 2015). Recent literature indicates that young people are more, less, or equivalently preoccupied with privacy. Studies focusing on the consequences of young people's privacy concerns similarly offer contradictory results, including both increased protective behaviours and the development of risky behaviours (Miltgen & Peyrat-Guillard, 2014). While a growing body of literature focus on privacy issues related to young people engagement with digital technology (Lusoli & Miltgen, 2009), comparative studies with adults are limited at best (Steijn & Vedder, 2015).

Furthermore, empirical research on offer typically uses convenience samples of college students, inadequate to address age effects. This leads to the question: does a generation gap exist? (Banks, et. al. 2014). Few studies based on nationally representative samples have examined the relationship between online privacy and age. Surveying German Internet users, Taddicken (2013) found that age had little relationship to SNS information disclosure, or privacy concerns. Similarly, based on representa-

tive US sample, Hoofnagle et al (2010) found no significant differences by age across a range of privacy variables. However two Pew surveys of representative samples of the US population both found that older users were less likely to have changed their privacy settings, deleted unwanted comments, removed their name from photos or taken steps to limit information about them on SNSs (Madden & Smith, 2010, Raine et al. 2013). In a recent UK study, Banks et al (2014) reports results that contradict some of the previous studies. In their survey almost 95% of 14-17-year-olds have checked or changed their privacy settings. From there the percentage who has taken action to protect their privacy drops to 32.5% of respondents age 65 and over. As the authors note, results from previous research are mixed, and the ambiguities surrounding age make it fertile ground for additional research. The present study aims to contribute further, providing up-to-date findings on the Israeli population.

3 Method

This research is based on a national survey recently conducted in Israel. Survey sampling unit comprised of Israeli households, individuals age 12+. Telephone interviews were performed by the Mahshov Public Opinion Research Company using a Scout system. Each sector was interviewed in its native language by trained surveyors with appropriate native language (Hebrew, Arab, Russian).

3.1 Sample

The sampling method involved random probability sampling of a proportional stratified sample. The sample included 1,052 subjects, sampled proportionally according to their percentage in the general population so that the sample size from each stratum (sector) is proportional to the size of the strata population. Maximum sampling error is plus or minus 3 points for the sample size with a 95% confidence interval. Sampling of sectors: veteran Jews—68%, immigrants who arrived in Israel after 1990—12%, Arab sector—20%. A proportional random strata sampling was also performed for each sector based on residence district. A real-time sample control was also performed based on the age and gender variables of each cell using the data published in the current annual reports of the Israeli Central Bureau of Statistics (CBS). A Kolmogorov-Smirnov test based on the CBS data, which was performed after the sampling, found that none of the distributions varied statistically from the CBS distribution apart from the Jewish sector sample. This sample had a statistically significant variance with a significance level of 1-5% for age distribution. The sampling design was planned as follows: controlled variables monitored district according to the CBS. Monitored variables monitored: gender, age, new immigrant (in order to preserve the limit for this population), religiosity (in order to monitor the ultra-orthodox – ‘haredi’ sector).

z-tests were conducted to establish the significance of distribution differences and a post-hoc Bonferroni test for multiple comparison was performed. Chi-square tests were also conducted in order to examine statistical independence between categorical variables. Throughout this paper, statistical significance $p < 0.05$ is indicated using coloured cells.

Within the nationally representative sample, age groups were assembled drawing on generational categorization most common in the literature (How and Strauss, 2000; Rosen, 2010). Specifically, we divided the sample as follows: iGeneration consisting of adolescents, with age ranging from 12 to 17 years of age. YGeneration included young adults aged 18-34; X Generation ranging from 35 to 54 and finally, the BabyBoom generation of adults aged 55 and over.

3.2 Measures

The survey instruments included measures to assess both behaviours and perceptions related to online privacy. In addition to general patterns of use, behavioural construct included self-disclosing behaviours and privacy-protective behaviours. Perceptual constructs consisted of privacy self-efficacy and privacy knowledge.

To measure self-disclosure behaviour, three items were developed to capture distinct disclosure activities common is the use of SNS namely, sharing photos, sharing location and tagging. Respondents were asked how often they engage in these activities.

Privacy protective behaviours were measured with items assessing two coping strategies: application removal and password management. The items were adopted from prior studies investigating privacy protection behaviours (cf. Moscardelli & Divine 2007) and revised in accordance with current technological affordances.

Privacy self-efficacy was measured using two items, which were developed by examining prior work on a related construct (LaRose, et al 2001). This study developed items estimating confidence (or lack thereof) in protecting privacy from e-services' information practices. The two items used in this study were: "I know exactly which tools to use in order to protect my privacy while using the internet" and "I feel powerless in the face of internet companies' ability to collect information about me while using the internet".

With regard to privacy knowledge, researchers have operationalized different dimensions associated with knowledge of data collection risk (Park, 2012). Building on previous studies, our measure for privacy knowledge was aimed at assessing the level of awareness to common institutional practices of data collection and use, particularly by commercial stakeholders. Three items were used in this study: "companies regularly document my habits while browsing and are utilizing my personal information for commercial purposes"; "I know exactly which personal information is being collected about me by internet companies, applications and websites"; and finally a false statement: "information I disclose or share online is never passed on to stakeholder whom I did not disclosed to directly".

4 Results

Results from the research pertain to key variables of behaviour and perception, as they relate to online privacy. In what follows we elaborate on the categories and their subsets, providing corresponding results.

4.1 Privacy-related behaviour

The general category of online behaviour was divided into two main subsets. First we explored use patterns, which we divided into *general use* and *self-disclosure behaviour*. Second, we examined *privacy protecting behaviour* more specifically.

4.1.1 Use patterns across generations of Israelis

Analysis of general internet use patterns across generations showed that 67% of Y generation goes online daily, several times a day. iGeneration followed with 65%. As age goes up the percentage of users reporting going online several times a day decreases with significant difference as indicated in table 1.

| | BB (55+) | X(35-54) | Y(18-34) | I (12-17) |
|--|----------|----------|----------|-----------|
| | 40% | 52% | 67% | 65% |

4.1.2 Use of social networking sites

Social networking sites (SNS) encourage users to share substantial amounts of information about themselves. Indeed, social networking behaviour is essentially based on self-disclosure (Chen, 2013). Survey results on the question: 'How often do you access SNS such as Facebook, Twitter, Waze etc.' is indicated in table 2 below, showing frequency of use between different age groups.

| | BB (55+) | X(35-54) | Y(18-34) | I (12-17) |
|-------------------------------|----------|----------|----------|-----------|
| daily, several time a day | 28% | 43% | 60% | 63% |
| daily, once a day | 17% | 11% | 14% | 14% |
| almost everyday | 6% | 7% | 8% | 10% |
| once or twice a week | 8% | 7% | 5% | 4% |
| once to several times a month | 4% | 3% | 0% | 1% |
| every 2-3 month or less | 2% | 1% | 1% | 0% |
| never use social networks | 35% | 28% | 12% | 8% |

In early 2014 claims were made regarding declining teen interest in SNSs, particularly Facebook (Kiss, 2014). Even though the present survey did not examine the frequency of Facebook use specifically but rather the frequency of SNS use as whole, results from Israel do not support the contention that teens are moving away from SNSs. Only 4% of adolescents aged 15-17 and young adults aged 18-24 said they had not registered on some kind of SNS. The percentage of high frequency internet users who use SNSs 'daily' or 'multiple times a day', is higher in the 12-34 age group and declines with age.

| | BB (55+) | X(35-54) | Y(18-34) | I (12-17) |
|--|----------|----------|----------|-----------|
| | 46% | 61% | 81% | 77% |

In sum, analysis of internet use patterns in general, and of SNS in particular, show significant differences across age groups, clearly dividing young generations (i and Y) from older generation (X and BB). As one would expect, young people are heavier users of digital technology and frequency of use decreases as age goes up.

4.1.3 Self-disclosure behaviour

Israeli internet users were asked whether they provide personal information such as ID number, home address, telephone number and other personal details when registering for eServices. 64% said they never provided such information. 6% said that they provided identifying information once but would not do so again. This compared with 28% who said that they did supply personal information when registering for various internet services. Generational analysis on this question revealed that significantly fewer adolescents (ages 12-17) supply personal information (17%) compared to users of older generations. This finding contrast with prevalent notions of risky online behaviour associated with adolescents. It corresponds, however, with recent studies showing lying behaviour displayed by youngsters as a privacy protection strategy ((Miltgen & Peyrat-Guillard, 2014).

| | BB (55+) | X(35-54) | Y(18-34) | I (12-17) |
|--|----------|----------|----------|-----------|
| | 32% | 32% | 22% | 17% |

4.1.3.1 Disclosure of personal information on SNSs

People who use SNSs routinely engage in self-disclosure practices online, which in turn affect their privacy. We examined three common practices and their frequencies: (1) sharing personal photos /videos on SNSs such as Facebook, Flickr, Picasa, and YouTube; (2) tagging personal photos, and (3) allowing such applications as Waze, Facebook, and Twitter to receive information about location. Results show variances in the percentage of users who engage in these practices. Half the SNS users (51%) shared personal photos or videos compared with 46% who did not share photos or shared them in the past but won't do so in future. 43% tagged themselves or others in photos versus 54% who did not tag or used to tag but would not do so in the future. 47% shared their location with application providers, compared with 50% who did not.

When these practices were tested for age, it was found that they mostly apply to younger users: as age go up the percentage of users enacting these practices declines. The percentage of Baby boomers (55+) that shared self-photos differed significantly from all other generations. A significant difference was found in photo tagging between the younger generations (i and Y) and the older generations (X and BB). Similarly, a significant difference was found in sharing geographical location between the under 35 users and the 35+ users. These results are summarized in table 5.

| | BB (55+) | X(35-54) | Y(18-34) | i (12-17) |
|--|----------|----------|----------|-----------|
| I share my photos/video | 28% | 48% | 56% | 63% |
| I tag myself / others in photos | 17% | 34% | 49% | 62% |
| I allow application to know my geographic location | 26% | 39% | 54% | 57% |

A direct significant correlation was found between the three practices examined and frequency of SNS use. When SNSs are used more frequently users will tend to share photos / videos, tag themselves, or let applications know their location. For example, roughly half the users (52%) who use SNSs multiple times a day tend to tag themselves in photos while only 32% of those who use SNSs once a day and only 16% of those who use SNSs twice a week or less, are likely to tag. A weak though significant correlation was found between the three practices and the frequency of SNS use, which persisted even after controlling for age and frequency of internet use. The heavy users who access SNSs multiple times a day differ significantly from other users in all three practices examined.

To conclude, young generations are heavier internet users in general and heavier disclosures, particularly in SNSs. Furthermore, analysis suggests that general use and self-disclosure are reinforcing one another, as use patterns and self-disclosure practices are positively correlated. Having addressed online self-disclosure behaviour, we move on to report results concerning privacy protective behaviour.

4.1.4 Privacy protective behaviour

Two kinds of online privacy protective activities were addressed. First we examined the extent to which users remove/uninstall applications due to demand for PII. Second, behavior related to password management was examined.

4.1.4.1 Application removal/uninstall

Users were asked whether they had removed / uninstalled applications or software from their PC or smartphone upon request for excessive personal information. 50% of teen users (iGen) and 52% of users aged 18-34 (yGen) uninstalled software/apps due to request for too much personal information. Such behavior decreases among older generations: 40% of users aged 35-54 (X Gen) and only one-third (32%) of users aged 55+ (BBs). The difference between the older and younger groups shows statistical significance.

| | BB (55+) | X (35-54) | Y (18-34) | i (12-17) |
|-------------------------------|----------|-----------|-----------|-----------|
| I don't download applications | 37% | 32% | 27% | 24% |
| Done so several times | 21% | 26% | 28% | 30% |
| Done so once | 11% | 14% | 24% | 20% |
| Never happened to me | 16% | 20% | 17% | 20% |
| I do not have PC/smartphone | 11% | 5% | 2% | 5% |
| I refuse to answer | 3% | 2% | 1% | 4% |

4.1.4.2 Changing Password for eServices

A security measure that can prevent unauthorized access to private online accounts is to change user password regularly. Nonetheless, only one third of all users (36%) change their passwords periodically. More than half (53%) never changed their password, and 7% changed their password in the past but had no plans to change it again. Analysis across generations shows no significant difference, as indicated in table 7 below.

| | BB (55+) | X(35-54) | Y(18-34) | i (12-17) |
|--|----------|----------|----------|-----------|
| I never changed my password | 53% | 53% | 52% | 56% |
| I change my password periodically | 39% | 34% | 36% | 36% |
| I've changed in the past but won't do it again | 5% | 7% | 9% | 4% |
| Not applicable/refuse to answer | 4% | 6% | 4% | 4% |

Concluding results on privacy protective behavior, we find that in terms of application/software removal due to demand for PII, young generations are significantly more cautious of their privacy compared with their older counterparts. As for password management, majority of all users never change passwords for their online accounts thus enacting privacy protective behavior, or rather lack thereof, in a similar manner.

4.2 Privacy Perceptions

We incorporated two key constructs to examine generational differences related to privacy perceptions: self-efficacy and knowledge. Self-efficacy was originally defined as one's belief in one's ability to succeed in specific situations (Bandura, 1977). This construct has been adopted in studies of privacy perception to capture and measure 'Privacy self-efficacy' that is, individuals' confidence in their abilities to protect their privacy from e-services' information collection and sharing activities (Youn, 2009).

4.2.1 Privacy Self-efficacy

4.2.1.1 Ability to use privacy protecting tools

Based on prior research, we developed two items estimating confidence (or lack of) in protecting privacy from e-services' information practices. These are presented next with corresponding results.

Digital platforms enable companies and other stakeholders such as friends, family, employers, and government agencies to gather information about individual users. However, some privacy protection tools are available online, at times free of charge. Are users familiar with such tools? Do they know which tools could potentially protect their privacy to certain extent? More than half of the respondents (57%) claim to know which tools to use. One quarter (25%) admitted that they did not know. We found a correlation between the rate of agreement with the statement 'I know exactly which tools to

use in order to protect my privacy while using the internet’ and age: as age goes up the rate of agreement with the statement decreases. Thus, almost three of every four adolescents (73%) claim to know which tools to use to protect their privacy, but as respondents' age increases, the percentage who agrees with the statement decreases and over the age of 55 only 42% of users agree they know which tools to use.

Table 8: “I know exactly which tools to use in order to protect my privacy while using the internet”

| | BB (55+) | X(35-54) | Y(18-34) | i (12-17) |
|--|----------|----------|----------|-----------|
| Agree | %42 | %50 | %62 | %73 |
| disagree | %33 | %29 | %22 | %12 |
| Sometimes agree and sometimes disagree | %13 | %15 | %14 | %10 |
| No opinion/not applicable | %12 | %6 | %2 | %5 |

4.2.1.2 Inability to control personal information online

Privacy self-efficacy was further measures with responses to the following statement: “I feel helpless in the face of internet company’s ability to collect information about me”- an emotional statement expressing feeling of frustration. 48% of the respondents agree with the statement, one third (33%) disagree, thus believing in one’s ability to fend off efforts to collect information about them. Here too, results suggest that young generations feel more capable to fend off information gathering techniques than adults. Nevertheless high proportions of all users share a feeling of helplessness in this regard.

Table 9: “I feel helpless in the face of internet company’s ability to collect information about me”

| | BB (55+) | X(35-54) | Y(18-34) | I (12-17) |
|--|----------|----------|----------|-----------|
| Agree | %50 | %53 | %44 | %44 |
| disagree | %28 | %29 | %38 | %36 |
| Sometimes agree and sometimes disagree | %12 | %12 | %14 | %13 |
| No opinion/not applicable | %10 | %6 | %3 | %6 |

4.2.2 Privacy knowledge

Building on previous studies, our measure for privacy knowledge aimed at assessing the level of awareness to common institutional practices of data collection and use, particularly by commercial stakeholders.

4.2.2.1 Knowledge of personal information collected by third party

Do Israeli internet users think they know which personal information is collected about them by services, software, and websites? Results are mixed. 40% say they do not know 'exactly' what personal information is being collected about them and 39% say they do know. Fewer users agree with this statement as respondents' age go up. For example, while 48% of iGeneration say they know exactly what information is collected about them, 37% of XGeneration agree and only 31% of BB aged 55+ agree with the statement. Generational differences are significant as indicated in table 10 below.

Table 10: I know exactly which personal information is being collected about me by internet services, applications and websites

| | BB (55+) | X(35-54) | Y(18-34) | i (12-17) |
|--|----------|----------|----------|-----------|
| Agree | %31 | %37 | %41 | %48 |
| disagree | %44 | %40 | %42 | %31 |
| Sometimes agree and sometimes disagree | %10 | %16 | %13 | %12 |
| No opinion/not applicable | %14 | %7 | %4 | %9 |

A major privacy issue associated with digital information sharing concerns use of PII by third party (Nissenbaum, 2010). We investigated users' awareness of such practice as function creep by asking their response to the following statement: "information I disclose or share online is never passed on to stakeholder whom I did not disclosed to directly". Analysis shows that nearly half (48%) the users disagree with the statement. In other words, they believe that at some stage personal information about them is transferred to other parties. About a third (35%) agreed with this statement. That is, they think the information they supply to a particular company remains with that company. 11% were neutral, 6% had no opinion. Analysis by age group reveals that as respondents' age increases so does the percentage of respondents realizing that information they disclose is transferred to other parties (leading them to disagree with the statement). A significant difference was found between iGeneration (30%) and all three older generation as indicated in table 11.

Table 11: information I disclose or share online is never passed on to stakeholder whom I did not disclosed to directly"

| | BB (55+) | X(35-54) | Y(18-34) | i (12-17) |
|--|----------|----------|----------|-----------|
| Agree | %26 | %34 | %36 | %46 |
| disagree | %55 | %48 | %48 | %36 |
| Sometimes agree and sometimes disagree | %8 | %13 | %13 | %9 |
| No opinion/not applicable | %10 | %5 | %3 | %9 |

4.2.2.2 Awareness of common surveillance practice

We asked respondents whether they agree that companies track their habits and use their data for commercial purposes. Results show that overall 54% agreed and 28% disagreed, while 12% provided the neutral response (sometimes agree, sometimes disagree), and 7% offered no opinion. We found significant difference between iGeneration and all three older generations as shown in table 12: youngster seems most naïve and less knowledgeable about common surveillance practice in the digital environment.

Table 12: companies regularly document my habits while browsing the internet and are utilizing my personal information for commercial purposes

| | BB (55+) | X(35-54) | Y(18-34) | i (12-17) |
|--|----------|----------|----------|-----------|
| Agree | %51 | %55 | %58 | %44 |
| disagree | %27 | %29 | %25 | %37 |
| Sometimes agree and sometimes disagree | %11 | %11 | %13 | %12 |
| No opinion/not applicable | %11 | %5 | %5 | %7 |

5 Discussion

This nationally representative study was designed to address dimensions of both behaviour and perception related to online privacy across four generations of Israelis. More specifically, we measured and compared behaviours involving self-disclosure and privacy protection. Privacy perceptions included two key construct, namely, privacy knowledge and privacy self-efficacy. In this short discussion section we reflect upon, and draw attention to, some of the key findings arising from this research.

As far as online self-disclosure is concerned this study comes to conclude that a digital generation gap exists between young generations of Israelis (i and Y generations) and older generations (X and BB). This conclusion finds support in a series of results, all of which pointing to significant differences in levels of digital participation and disclosure behaviours. Reflecting on the nature of this apparent gap, we note that contrary the previous findings (cf. Hoofnagel et al, 2010) young users seem conscious of their disclosure and of the associated privacy risks. This becomes evident from patterns of privacy protective behaviour as discussed further on. Similar findings also emerged from a recent European-based study (Miltigen & Peyrat-Guillard, 2014). We note, however, that self-disclosure practices among younger generations does not appear entirely optional. Indeed, over 90% of this cohort enacts such practices compared to 75% (XGen) and 52% (BB) as age goes up. We propose that young users are compelled to take privacy risks because digital participation, which necessitates disclosure, is a social must. In contrast, older users are seen to disclose less and with more discretion. For generation X and more so, for Babayboomers, digital abstinence or partial participation still remains a practical option. We therefore concur with Blank et al (2014) contention that digital technology have become so embedded in the social lives of users that to maintain their social lives they must disclose information on them despite the fact that there is a significant privacy risk in disclosing this information. Hence, at the heart of the generation gap lies not the issue of awareness as previously suggested (Barnes, 2006) but rather the social price and lost opportunities (Livingstone, 2008) of preserving privacy by restraining self-disclosure - an intolerable price for the digital natives, yet an acceptable one for the ‘immigrants’.

On the issue of privacy protective behaviour however, a different picture comes into view. Contrary to widespread impression, our findings show that young internet users do not do less to protect their

online privacy compared with their older counterparts. If anything, they seem to be doing more. To the extent a generational gap exist at all – it operates against the commonly held view of young generations as lacking concern for privacy and by implication, doing less to protect it. For example, our finding that young generations uninstall software / applications due to demand for PII significantly more compared with older generation may well indicate awareness to privacy risk otherwise why take measures to mitigate it? The few studies that have addressed this issue using representative samples (in UK, USA, and Australia) arrived at a similar conclusion, namely, that young people are much more likely to have taken action to protect their privacy online (Blank et al, 2014). Furthermore, findings from recent ethnographic research reveal the ways in which young people navigate privacy online. Bypassing prescribed technological affordances such as privacy settings, ignoring Terms of Service, and enacting a range of subversive practices (Raynes-Goldie, 2010; Young & Quan-Tasse, 2013) such as deactivating accounts, manipulating access to meaning (Marwick & Boyd, 2014) and providing false information (Miltigen, & Peyrat-Guillard, 2014) are examples of the strategies observed. We speculate that such obfuscation practices (Brunton, & Nissenbaum, 2015) are less common with adults. This may also account for findings regarding privacy self-efficacy showing significant generational differences in perceived ability to protect one's privacy online. We found that 73% of iGens believe in their ability to use privacy protecting tools compared to 62% of Ygeneration, 50% Xgenerations and only 42% within BBs. Additionally, we found that a sense of *helplessness* to control personal information online correlates with age. Although statistical significance was found between age groups, we note the large proportion of the population - nearly half - feeling powerless in the face of internet services' ability to collect and use information about them (44% of both iGeneration and Ygeneration; 53% generation X, and 50% BB).

Drawing on these findings, we argue that privacy solutions in the digital environment can no longer focus on the sole responsibility of individual users as implied by the privacy self-management paradigm currently dominating the internet (Solove, 2013). Our results regarding privacy knowledge suggest naivety (Hoofnagle et al 2010) is still at play, particularly within iGeneration but among older generations as well, suggesting remedy in the form of education and awareness. However, users are increasingly helpless and seem gradually to become aware of it. As online privacy becomes ever more networked (Marwick and Boyd, 2014) technological means of control based on individual notions of privacy becomes obsolete.

6 Conclusion

The result of the exploratory survey analysed in this paper reveals the intricacies involved with online privacy behaviour and perception of Israeli internet users representing four generations. Contrary to simplistic notions of a radical generation break, and of privacy perceptions clearly dividing the 'old who cares' vs. the 'young who doesn't care', analysis portrays a more nuanced picture. While a generational gap does exist in certain dimensions of privacy related behaviour and perception, it does not concur with widespread impressions and popular media myth. Young generations are heavily involved online; as such, they enact more practices of self-disclosure through a variety of social media affordances, and with greater frequency. However, the enactment of risky behaviour is both conscious and inevitable. Findings do not support the view of young generations as lacking concern for privacy since clearly they do more rather than less to protect it, as compared to their older counterparts. Debating as we are on the future of online privacy policies (Aquisti et al, 2015), arguments drawing on the allegedly reduced appreciation of privacy by young generations can have important implications for the development of future policies (Steijn & Vedder, 2015). Further research is called for, to increase understanding of individuals' appreciation of privacy and contribute to informed decision making better addressing the needs and concern of current and future generations.

7 Acknowledgments

The authors would like to thank Saar Gershon from the Hebrew University in Jerusalem, for his insightful statistical analysis.

The authors would also like to thank Microsoft Israel for their research grant that supported this study.

8 References

- Acquisti, A., Brandimarte, L., George, A. (2015) Privacy and Human Behavior in the Age of Information. Science.
- Antón, A., Earp, JB, & Young, JD.(2010) How Internet Users' Privacy Concerns Have Evolved since 2002. IEEE Security & Privacy
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84 pp 191–215.
- Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday* 10(3), 393–411.
- Blank, G., Bolsover, G. & Dubois, E. (2014) The New Privacy Paradox: Young people and privacy on social network sites Global Cyber Security Capacity Centre: Working Paper
- Boyd, D. (2014). *It's Complicated: The Social Lives of Networked Teens*. New Haven: Yale University Press.
- Brunton, F. & Nissenbaum, H. (2015) *Obfuscation: A User's Guide for Privacy and Protest* Cambridge: MIT Press.
- Chen, R. (2013) Living a Private Life in Public Social Networks: An Exploration of Member Self-Disclosure, *Decision Support Systems*, 55(3), pp. 661-668
- Dinev, T (2014) Why would we care about privacy? *European Journal of Information Systems* (2014) 23, 97-102
- Herring, S. C. (2008) Questioning the Generational Divide: Technological Exoticism and Adult Constructions of Online Youth Identity. In D. Buckingham (Ed.) *Youth, Identity, and Digital Media*. Cambridge, MA: The MIT Press, pp. 71-92
- Hong, W. & Thong, J. (2013) Internet Privacy Concerns: an integrated conceptualization and four empirical studies. *MIS Quarterly* 37 (1) pp. 275-298
- Hoofnagle, C. J., King, J., Li, S., & Turow, J. (2010). How Different are Young Adults from Older Adults When it Comes to Information Privacy Attitudes and Policies? (SSRN Scholarly) Paper No. ID 1589864). Rochester, NY: Social Science Research Network.
- Howe, N., Strauss, W. (2000). *Millennials Rising: The Next Great Generation*. New York: Vintage Books
- Kiss, J. (2014). Teenagers migrate from Facebook as parents send unto them friend requests. *The Guardian*. <http://www.theguardian.com/technology/2013/dec/27/facebook-deadand-buried-to-teens-research-finds>
- Livingstone, S (2008) Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression. *New Media & Society* 10 ,3 393-411

Lusoli, W. & Miltigen, C. (2009) Young people and emerging digital services: an exploratory survey on motivations, perceptions and acceptance of risks. JRC Scientific and Technical Reports EUR 23765 EN, EC JRC IPTS, Sevilla.

Madden, M., & Smith, A. (2010). Reputation management and social media. PEW Research http://www.pewinternet.org/~media/Files/Reports/2013/PIP_AnonymityOnline_09051

Marwick, A. & boyd, d. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society* (online)

Miltigen, CA, & Peyrat-Guillard, D. (2014) Cultural and generational influences on privacy concerns: a qualitative study in seven European Countries. *European Journal of Information Systems* 23, 103–125

Moscardelli, D & Divine, R. (2007) Adolescents' Concern for Privacy When Using the Internet: An Empirical Analysis of Predictors and Relationships With Privacy-Protecting Behaviors. *Family and Consumer Sciences Research Journal* 35(3)232-252

Nissenbaum, H (2010) *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, CA: Stanford University Press

Nussbaum, E (2007) Say Everything. <http://nymag.com/news/features/27341/>

Pavlou, PA. (2011) State of the information privacy literature: where are we now and where should we go? *MIS Quarterly* 35(4), 977–988.

Prensky, M. (2001) Digital Natives, Digital Immigrants. *On the Horizon*, 9 (5) 1-6

Rainie, L., Kiesler, S., Kang, R., & Madden, H. (2013). Anonymity, Privacy, and Security Online. http://www.pewinternet.org/~media/Files/Reports/2013/PIP_AnonymityOnline_09051

Raynes-Goldie, K. (2010) Aliases, creeping, and wall cleaning: Understanding privacy in the age of Facebook *First Monday*, 15 (1)

Rosen, L.D. (2010) *Rewired: Understanding the iGeneration and the Way They Learn*. Palgrave Macmillan

Steijn, W. & Vedder, A (2015) Privacy under Construction: A Developmental Perspective on Privacy Perception. *Science, Technology, & Human Values* 40 (4) pp. 615-637

Solove, D J.(2013) Privacy Self-Management and the Consent Dilemma. 126 *Harvard Law Review* 1880 SSRN: <http://ssrn.com/abstract=2171018>

Taddicken, M. (2013). The “Privacy Paradox” in the Social Web: The Impact of Privacy Concerns, Individual Characteristics, and the Perceived Social Relevance on Different Forms of Self-Disclosure *Journal of Computer-Mediated Communication* 19 (2) pages 248–273

Tufecky, Z. (2012) Facebook, Youth and Privacy in Networked Publics. <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/viewFile/4668%26lt%3B/5001>

White, D. and Le Cornu, A. (2011) Visitors and Residents: A new typology for online engagement *First Monday*, 16(9 - 5)

Youn, S. (2009) Determinants of online privacy concern and its influence on privacy protection behaviours among young adolescents. *Journal of Consumer Affairs* 43(3), 389–418.

Young, A. & Quan-Haase, A. (2013) privacy protection strategies on Facebook, *Information, Communication & Society*, 16(4)

Wang, Q., Myers, M. D., & Sundaram, D. (2013). Digital Natives and Digital Immigrants. *Business & Information Systems Engineering*, 5 (6), 409-419

RFID ENABLED VISUALIZATION OF PRODUCT FLOWS: A DATA ANALYTICS APPROACH

Anastasia, Griva¹, Department of Management Science and Technology, Athens University of Economics and Business, Athens Greece, an.griva@aueb.gr

Cleopatra, Bardaki¹, Department of Management Science and Technology, Athens University of Economics and Business, Athens Greece, cleobar@aueb.gr

Katerina, Pramataris¹, Department of Management Science and Technology, Athens University of Economics and Business, Athens Greece, k.pramatari@aueb.gr

Complete Research

Abstract

Radio frequency identification (RFID) is an information facilitator that can directly improve decision-making; thus many retailers and suppliers have adopted it. A vast amount of RFID data streams has been gathered, however, it remains unutilized, or it is been exploited solely for inventory count purposes. This research proposes a way to analyze the immense volume of RFID data reflecting the behavior of products in retail stores, in order to produce information for inventory availability and inventory flows at different stages of the supply chain. We propose an RFID data analytics artifact that transforms RFID data captured in retail stores to the flows of the inventory/ products between locations in the stores. By mining the RFID data streams, we reveal the flow patterns of the products; these patterns correspond to the frequent product paths in the stores, and we provide them to the retailers in a visual manner. This unprecedented knowledge is valuable, because it can enable decisions ranging from shelves space allocation, dynamic pricing programs for slow-moving fresh products to product assortment. Furthermore, to testify artifacts' correctness and usefulness, we have put it in practice, using real data provided by an Italian fashion retailer, in order to show how it can really support such decisions.

Keywords: RFID data analytics, decision making, product flow patterns, fashion retail

¹ ELTRUN: The e-Business Research Center

1 Introduction

Radio frequency identification (RFID) assists retailers to look into the operation of their business processes, and appears to be an important generator of innovative business information that has led to various organizational benefits. RFID spurs lots of prospects for efficient management of supply chain processes and decision support (Dimakopoulou et al., 2014). Thus, RFID technology should be viewed as an information facilitator that can directly improve decision-making (Sellito et.al, 2007). It plays an important role in supporting logistics and supply chain processes. It can provide suppliers, manufacturers, distributors and retailers precise real time information about the products (Zhu et al., 2012). Over the past two decades, the retail industry has implemented RFID resulting in reported improvements in inventory productivity (Unger & Sain, 2015). RFID technology could be also utilized to improve product traceability and the visibility, however, many retailers and suppliers/manufacturers have applied RFID solely for inventory management and inventory accuracy purposes (Hardgrave et. al, 2011, 2013; Thiesse & Buckel, 2015; Tijun et. al, 2015).

At the same time, business analytic techniques have been developed that can connect large datasets to enable broader and deeper analysis than previously possible (Provost & Fawcett, 2013; Phan & Vogel, 2010). As a consequence, data-driven decision making is now recognized broadly, and there is growing enthusiasm for the notion of "Big Data". Big Data analytics now drives near every aspect of our modern society, including retail industry, financial services etc. (Bertino et.al, 2011). Big Data research looks at how to analyze data in different domains with such characteristics; and in a way that generates deeper knowledge and adds value to the decision making process in businesses (Sharda, Asamoah, & Ponna, 2013).

Taking advantage of the new technologies, which allows us to process the immense volume of the RFID data streams, this research proposes how we can analyze the available vast amount of RFID data, reflecting the behavior of products in the retail supply chain, in order to produce information for inventory availability at different stages of the supply chain. We propose an RFID data analytics artifact that transforms RFID data captured in retail stores to the flows of the inventory/ products between locations in the stores. We mine the flow patterns; the frequent paths of products and provide them to the retailers/ suppliers in a visual manner. This unprecedented knowledge is valuable, because it can enable decisions such as, shelves space allocation to product categories, according to their moving rate. We have put the artifact in practice, using real data provided by an Italian fashion retailer, in order to show how it can really support such decisions.

The remainder of the paper is organized as follows. Section 2 describes the proposed methodology. Next, section 3 presents the application of our approach to real item level RFID data. Then, we highlight the decisions that could be supported by the mined knowledge. Finally, we overview this research's main outcomes; and discuss further research.

2 RFID Data Analytics Approach

2.1 Overview

This research has adopted the "Design Science" approach (Hevner et al. 2004). An artifact has been developed; it is an RFID Data Analysis Pipeline which consists of four phases. The proposed approach has been evaluated in practice; we used 6 month RFID data from a store of an Italian fashion retailer, to mine useful knowledge. Ultimately, the findings from the RFID data analysis aspire to give decision support to the Italian retailer and, generally, the apparel and fashion retail industry. Figure 1 depicts the proposed approach consisting of 4 phases: (A) Data Description & Preparation, (B) Modeling Product Flows, (C) Knowledge Mining, and (D) Business & Data Understanding. Business and data understanding phase is a "superset" of the other three phases; we need to apply it throughout the data analysis pipeline because the business goals need to be taken in account during the analysis, as well as the data derived from each step. Each phase has inputs and outputs; by refining the inputs of each

phase, and passing through the entire pipeline’s lens, we will extract the valuable knowledge which will be used to support decision making.

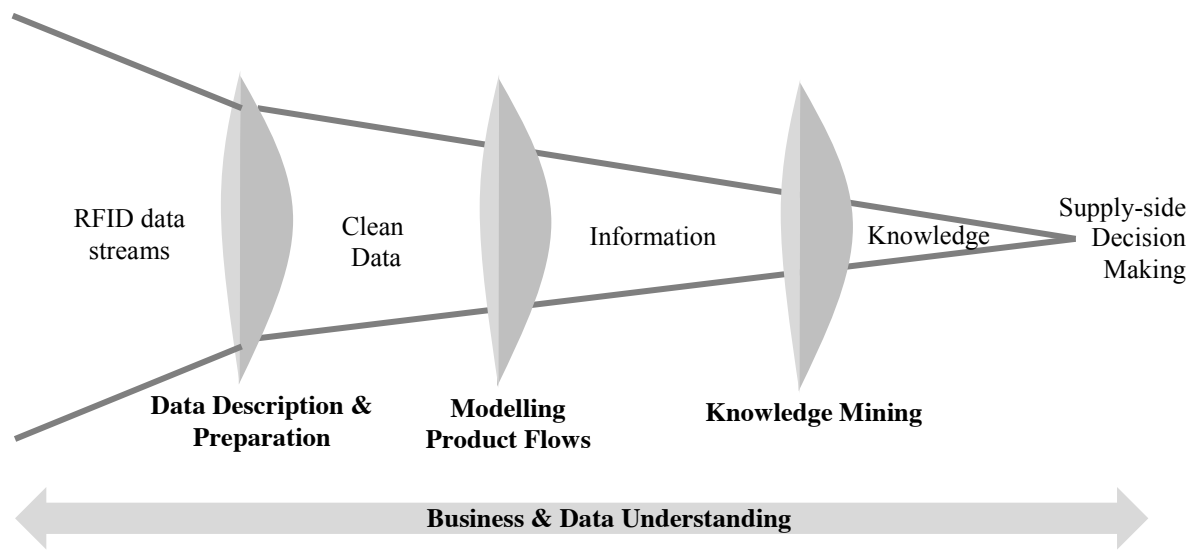


Figure 1. RFID Data Analytics Approach

2.2 Data Description and Preparation

The initial phase focuses on obtaining and exploring the dataset, having in mind the business objective, which is to extract knowledge to support decision making. The dataset needed to perform this analysis is RFID data that represent garments movements into the stores, made either by the customers or by the clerks while performing business processes and RFID-enabled processes such as replenishments, receiving, purchasing etc. We have to know the exact places of the products and the processes that are connected with these places (table 1). For example, we should know that a product, which is being uniquely identified by an electronic product code (EPC), has just been received in a specific time and it’s in the backroom, or the product has been transferred from backroom to store floor, and then it is located in check-out desk as it is sold, and at the end it passes the electronic article surveillance (EAS) gate. Moreover, it could be useful to have additional characteristics of the stock keeping units (SKUs), for example if the RFID captions concern garments we could have garments’ size, color, the product category they belong to, or in other cases we could have date of expire, products dimensions etc.

| EPC | Location | Process | Time Stamp | SKU | Additional characteristics |
|-----|--------------------|---------------|----------------|----------------|----------------------------|
| xyz | backroom | receiving | 4/7/2015 11:00 | 12443234323467 | ... |
| xyz | replenishment gate | replenishment | 5/7/2015 9:30 | 12443234323467 | |

Table 1. Dataset Example

The dataset could be enriched with data captured by real time location RFID system (RTLS) i.e. exact store areas in the store floor. For example, a product X in time Y is in store floor in “A” store area, and then in time X+1 a customer took the product and transferred it to “B” store area, and then the same product entered the fitting room also with product Z. The last dataset could enlighten the knowledge extracted from the analysis and could reveal the whole benefits of adopting RFID in retail and applying data analytics in the derived dataset.

Data preparation covers all the activities to construct the final dataset. Hence, we have to select the appropriate RFID data that will be used in the forthcoming analysis, and clean the dataset. As it is well-known RFID data streams have a noisy nature, thus we have to de-noising them by eliminating duplicate values, inconsistencies from the dataset, and corrupted values (Keller, et al, 2014).

2.3 Modelling Product Flows

A modeling precondition is to identify model's entities. This task is critical as it will help us to have a holistic understanding of the RFID system and the RFID-captured product movements. Thus, we have to identify the:

- Store locations the product movements take place. For example, backroom, sales floor, and their sub-locations. For example, the check-out, the replenishment gate, the EAS gate, sub-sectors in the store floor location etc.
- RFID readers. For example, RFID cages, replenishment gates, check-out readers, EAS gates, handheld readers etc.
- Business processes and RFID-enabled processes. For example, inventory, replenishment, check out, EAS gate, return, theft etc.

According to the RFID locations, the readers, and the processes take place in the store we have to identify the RFID captured product flows. It could be useful to create a graph in order to have a better visualization of these flows; this could be also used as input to predict garments' movements. In this graph we could depict the correct captured flows from a process and a system perspective, and the false product flows which lead to missing reads. This information could be used to detect the data quality issues, which derived from the readers' inconsistencies; for example, if a product is being captured in the receiving, and then directly in the check-out, we assume that there is a missing read in the replenishment gate.

Concerning the data quality issues, it's critical to identify the false positive self joins. The self joins are more than one contiguous in time captures from the same reader. In the dataset, we may observe that the same EPC is read more than one time respectively from a reader during passing from a location to another. For example, a garment in the process "store floor replenishment" is read two times respectively. This seems to be right from a system and a process perspective, as it is legit the reader to capture the garment two times in a row within few seconds. But in these self-join cases maybe there are hidden false positive self joins. These joins concern wrong RFID reads from a process perspective. These are self-join reads which happen in greater contiguous timestamps, and lead to missing RFID reads. For example, a product which is read two times when passing from backroom to store floor, with a time difference of two days, this is not a right self-join read, since other missing RFID reads that intervene.

2.4 Knowledge Mining

Finally, we have to transform RFID data into decision making data, according to the business questions we want to answer. Indicatively we could:

- Calculate the mean times between the basic entities of the model, and the entities of the transition graph, in product categories, SKU level, colors, sizes etc.
- Calculate the product categories, SKU, colors, and sizes in fast, medium, slow, and no moving, according to the each product's speed history on the sales floor. By history speed we refer to the average time a product spends on the sales floor from the time it was moved there till the time it was sold
- Classify the product categories, SKUs, color, sizes in high, medium, low and no sellers according to their sales.
- Detect the missing SKUs from the store that are high sellers and fast movers.

- Detect the missing sizes and colors of a high seller category.
- Detect of the garments that frequently enter the fitting room but they aren't sold
- Detect the misplaced products

3 Experiment

In this section we apply the above method in real data provided by a fashion retailer, in the context of SERAMIS² EU project, in order to testify the method's effectiveness.

3.1 Data Description and Preparation

The dataset will be analyzed concerns garment's movements captured via RFID in a retail store collected from 27 May to 30 November of 2014. The first step of this phase is to select the data that will be used for the forthcoming analysis. We decided to use the information about store locations, RFID readers, RFID processes, garment sizes, names, and colors, as shown in table 2. Then, we had to de-noise the dataset. So, by eliminating duplicates, corrupted values and inconsistencies, from the 307.243 initial record we continued the analysis with 99,89% (306.912 records) of them.

| EPC | Location | Process | Readpoint | Time Stamp | SKU | Category | Product Name | Size | Color |
|--|------------|---------------|--------------------|----------------------------|----------------|----------|--------------|------|-------|
| urn:epc:tag:gid-96:31161137.20037.100004 | checkout | Check out | checkout | 2014-10-12 17:06:20.000 | 31161137020037 | shirt | MOGOL | 48 | 3 |
| urn:epc:tag:gid-96:31161137.20037.100004 | storefloor | Replenishment | replenishment gate | 2014-09-17 16:26:09.000 | 31161137020037 | shirt | MOGOL | 48 | 3 |
| urn:epc:tag:gid-96:31161137.20037.100004 | backroom | Receiving | RFID cage | 2014-09-04 11:17:15.000 | 31161137020037 | shirt | MOGOL | 48 | 3 |
| urn:epc:tag:gid-96:31161137.20037.100004 | storefloor | Inventory | handheld1 | 2014-09-23 15:33:33.000 | 31161137020037 | shirt | MOGOL | 48 | 3 |
| urn:epc:tag:gid-96:31161137.20037.100004 | storefloor | Inventory | handheld1 | 2014-09-30 16:07:23.000 | 31161137020037 | shirt | MOGOL | 48 | 3 |
| urn:epc:tag:gid-96:31161137.20037.100004 | storefloor | Inventory | handheld1 | 2014-10-06 12:01:59.000 | 31161137020037 | shirt | MOGOL | 48 | 3 |
| urn:epc:tag:gid-96:10160339.60054.100003 | backroom | Replenishment | replenishment gate | 2014-08-18 13:30:32.000 | 10160339060054 | cloak | LISA | 42 | 5 |

Table 2. Sample of the selected RFID data

3.2 Modelling Product Flows

Our first goal is to depict the RFID streams in a graph, and display garments RFID-captured movements. Identifying the model's entities as shown in figure 2, we have (A) the store locations (B) the RFID readers, and (C) the RFID processes. Concerning the store locations, we have the backroom, and the backroom entrance from where the garments enter the store, the store floor, the replenishment gate, which is the intermediate location between backroom and store floor, the store aisles, where the garments are displayed, and the check-out desk where the garments are moved during the purchasing process. Each store location has RFID readers. We have two kinds of readers, the stationary and the non-stationary readers. The last of them could be used either in backroom or in store floor for inventory control. Moreover, we have 7 RFID processes; each process is connected with a reader. In the next paragraphs we will eliminate those processes that interact with the "out store environment" i.e. C5, C6, C7. Last but not least, at this point we have to mention that the inventory processes may happen by clerks at any time via using handhelds. In figure 3 are shown the entities of the model in the store layout.

² Sensor-Enabled Real-World Awareness for Management Information System (<http://seramis-project.eu>)

- A. Store locations:
1. Backroom
 - i. Backroom Entrance \ni Backroom
 2. Replenishment Gate (an intermediate store location)
 3. Store floor
 - i. Checkout Desk \ni Store Floor
 - ii. Store Aisles \ni Store Floor
- B. RFID Readers:
1. Stationary:
 - i. RFID Cage (is placed in location A1i)
 - ii. Replenishment Gate (A2)
 - iii. Checkout reader (A3i)
 2. Non-stationary:
 - i. Handhelds: (A1,A3ii)
- C. RFID Processes:
1. Receiving (B1i)
 2. Replenishment (B1ii)
 3. Inventory (B2i)
 4. Checkout (B1iii)
 5. No RFID in Arrival (B1i)
 6. Shipping to Distribution center (B1i)
 7. Shipping to other Stores (B1i)

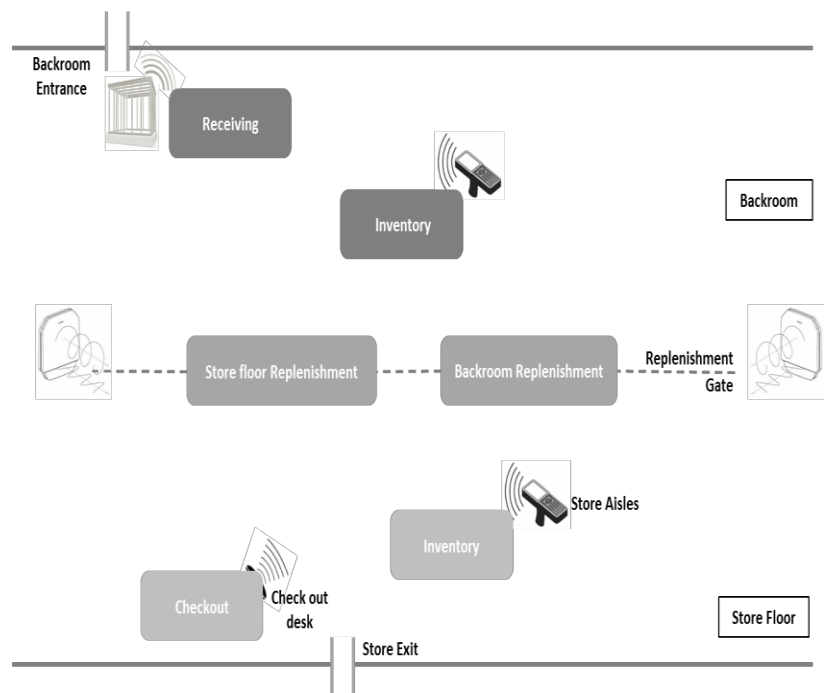


Figure 2. Basic entities of the model-graph Figure 3. Model's entities in the store's layout

According to the recognized model's entities, we identify the RFID-captured garments' movements, and we create the garments' flow graphs. In figure 4 is shown the garment's graph which concerns the correct flows. The number in each node represents the RFID captured garments at the location. The green arrows represent the garments' correct flows. The flows from a node to the end, show the number of garments which passed from this node and stayed there, so during the time frame of the given dataset we can't detect them in any other node after that event. For example, at figure 4 we can notice that the majority of the garments (70.211) had been captured for first time at the backroom entrance during the receiving process via the RFID cage. At about 35% of them (24.125) stayed at the backroom, and had never been displayed at the store floor. We have only eleven captures in the backroom during the inventory count process, via the handheld. At about 42.852 garments passed from the backroom to the store floor through the replenishment gate, and 27.720 garments had been sold.

As far as the self-loops of the replenishment transitions are concerned, we came to the conclusion, that they are also divided into right flows, and false positive flows. The right self-loops are those that the time from the first timestamp (t_1) until the next one (t_{1+1}) is less than one minute. This means that the garment stayed in this location and by mistake it had been read more than once. The above also happens for the RFID cage self-loops. We assume that the other self-loops at the replenishment gate are wrong. This happens because if $t_{1+1} - t_1$ is more than one minute maybe between these periods of time there are missing intermediate reads. Concerning the self-loops of the inventory nodes, we accept that all of them are right, because as mentioned before, it's possible to check the inventory a lot of times during a period, before the garment reach to change a node. Last but not least, in each transition graph it could be also meaningful to convert the actual number of garments pass from each node into percentages, in order to show the probability of which will be the next node of the garment the moment $t+1$.

In the same context, we also depict in figure 5 the garments' false flows. By examining the wrong flows we could figure out inconsistencies in the RFID readers. Thus, we could improve the quality of the data derived from the RFID system.

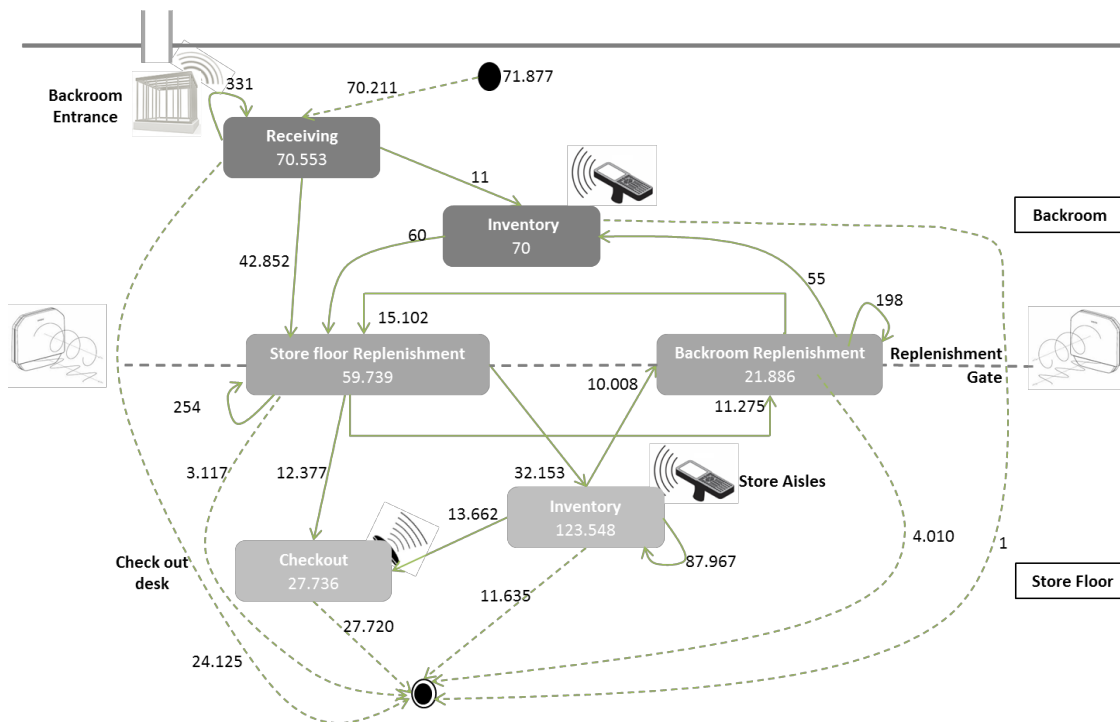


Figure 4. Garments Transitions Graph- right flows

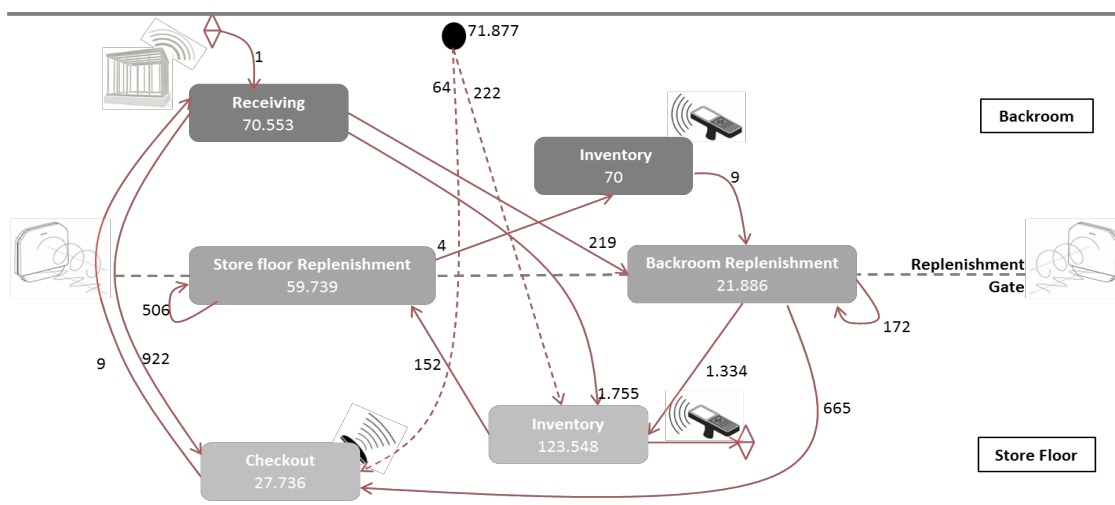


Figure 5. Garments Transitions Graph- wrong flows -> missing reads

3.3 Knowledge Mining

After eliminating from the dataset the false flows we need to transform the RFID data into decision making data. In this section we will show what knowledge we could mine from the given dataset, which business questions we could answer, and what business reports we could support. From the data provided we could extract the average days the garments stay in each location before being transferred to another. For example, in figure 6, we could figure out that 27.068 garments stayed on average 2 months (60,49 days) in the store until they have been sold. This report could be useful for the store manager in order to observe garments states, and detect delays. Moreover, it would be more meaningful to have this kind of report per product category, or per color, size and SKU, and this question could be answered from the given dataset.

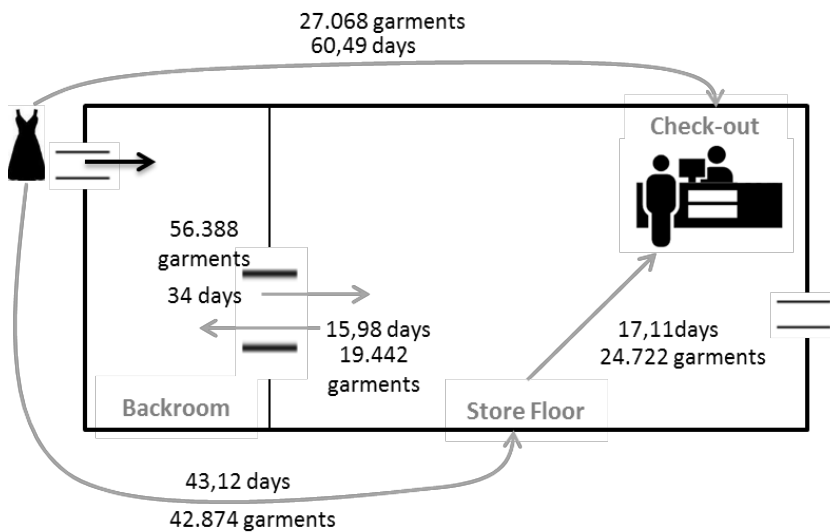


Figure 6. Average days per location

Furthermore, we could categorize the garments in fast, medium, slow, and no moving, according to the average time from the moment they being transferred from the backroom to the store floor, till they have been sold. This could be helpful for the store’s manager to determine replenishment strategies. According to five-number summary (box plot) as shown in figure 7 the garments that stayed from 4,33 to 10,88 days in the store floor before they been sold are considered as fast moving etc. Then utilizing the box plot we created figure 9 that depicts the product categories in classes according to how fast they are bought by customers. The green product categories are the high moving ones, the orange are the medium moving ones, and the red are the slow moving ones. At this point we have to mention that before calculating the five-number summary the outliers had been eliminated. Outliers are product categories that moved from store floor to check out only a few times (i.e. less than 20). This report is useful to be combined with the top selling product categories (figure 8). This way the manager will observe and compare the fast moving and the top selling categories. This information could be utilized, for example, to decide the space allotted for each category, the number of items per color or per size that should be displayed in the store floor. Moreover, the categories which are unsold in the store floor are labeled as “no moving”; in our case these categories are “leather skirt” and “short trousers”. The following figures concern the results extracted by all the given timeframe (6 months), thus it could be also useful to have this kind of information daily, in order to have a better visibility of the garments.

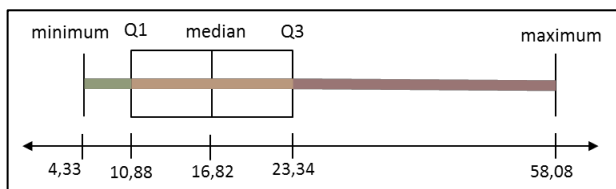


Figure 7. Product Categories – Box Plot in average time from store floor to check out

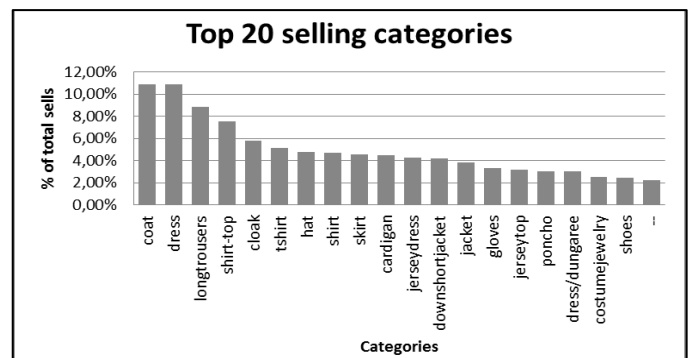


Figure 8. Top selling product categories

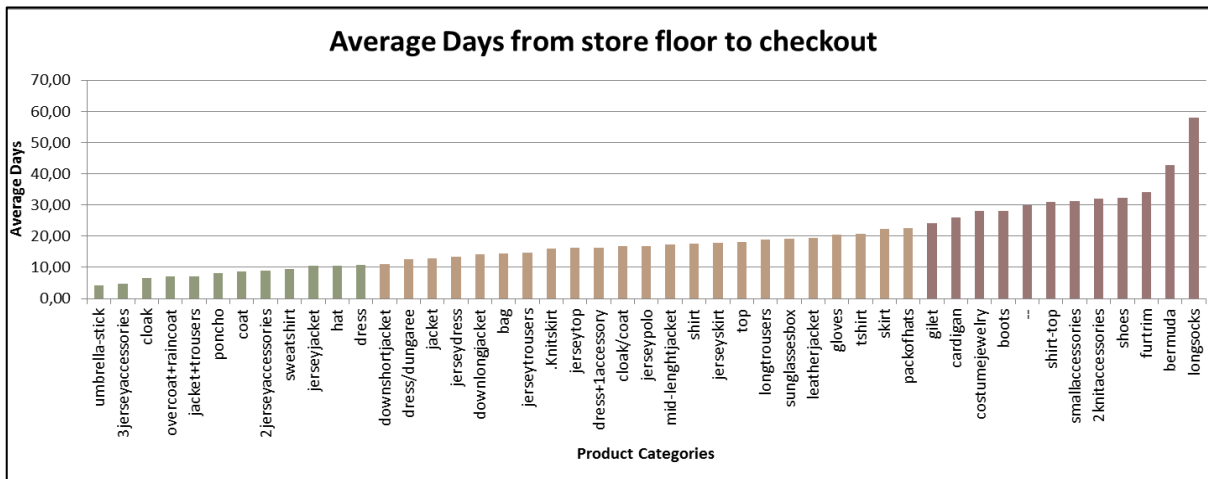


Figure 9. Product categories classification based on average days from store to check out

The same reports as above could be also generated per sizes and per colors. For the SKU level it could be more useful to produce a table with the fast moving SKUs compared with the top SKU sellers, as shown in tables 3 and 4. The timespan used for these descriptive is all the available dataset i.e. from 27 May to 30 November of 2014.

| SKU | Category | Product Name | Size | Color | Average Days to be sold |
|----------------|----------------|--------------|------|-------|-------------------------|
| 63740141050601 | poncho | NANO | NULL | 60 | 0,13 |
| 63740141050031 | poncho | NANO | NULL | 3 | 0,60 |
| 63740041050601 | poncho | NAFTA | NULL | 60 | 1,26 |
| 63740041050151 | poncho | NAFTA | NULL | 61 | 1,60 |
| 83740643050421 | poncho | EMILIO | NULL | 42 | 3,34 |
| 65140343050601 | bag | KITTI | 36 | 60 | 4,33 |
| 65340132050011 | umbrella-stick | MAN | 36 | 1 | 4,33 |

Table 3. Fast moving SKUs

| SKU | Category | Product Name | Size | Color | No of Garments |
|----------------|------------------|--------------|------|-------|----------------|
| 85749840050011 | hat | CAPP-AI | 36 | 1 | 1027 |
| 85749830050011 | hat | CAPP-AI | 36 | 1 | 108 |
| 65340132050011 | umbrella-stick | MAN | 36 | 1 | 41 |
| 57560838060012 | costume_Jewelery | AMICO | NULL | 1 | 35 |
| 63740041050151 | poncho | NAFTA | NULL | 15 | 34 |
| 63740041050601 | poncho | NAFTA | NULL | 60 | 33 |
| 83740343050421 | poncho | EROS | NULL | 42 | 32 |

Table 4. Top Selling SKUs

Furthermore, it would be helpful for the manager to have a list with the SKUs, from which they have a lot of pieces, but are kept in the back store without having a presence in the sales floor, as shown in table 5 from the 6 month data we analyzed.

| SKU | Category | Product Name | Size | Color | No of Garments |
|----------------|------------------|--------------|------|-------|----------------|
| 26060629020012 | various | PEPE | NULL | 1 | 20 |
| 48040428030041 | sunglasses | CALYPSO | NULL | 4 | 17 |
| 45260224060011 | shoes | IRIVOLO | 39 | 1 | 16 |
| 27240527030013 | smallaccessories | SESTRI | NULL | 1 | 15 |
| 47540646050281 | costume_Jewelery | CRI8 | 36 | 28 | 15 |
| 48040128030023 | sunglasses | CALESSE | NULL | 2 | 15 |

Table 5. SKUs that remain in the backroom without having a presence on sales floor

Moreover, another interesting report would be to present the “no moving” SKUS that are presented on the sales floor, but are for a long period unsold. For example, for the 6 month data in table 6 are shown SKUs that stayed from 4,5 to 5 months unsold on the sales floor.

| SKU | Category | Product Name | Size | Color | Average Days Unsold |
|----------------|------------------|--------------|------|-------|---------------------|
| 53460139060014 | cardigan | GHIANDA | NULL | 1 | 160,08 |
| 90140136030013 | cloak | LISA | 40 | 1 | 159,65 |
| 60140430030027 | cloak | DISCANTO | 48 | 2 | 159,30 |
| 90140136030014 | cloak | LISA | 42 | 1 | 156,25 |
| 47241037030013 | smallaccessories | AMMIRARE | NULL | 1 | 154,88 |
| 90140136030012 | cloak | LISA | 38 | 1 | 149,28 |

Table 6. SKUs that are present at the sales floor for a long time unsold

The above could be utilized in cases with little space on the sales floor in order to achieve a better allocation of the available sales area. Moreover, in these cases if we don't have “no moving” SKUS, we could move “slow moving” SKUS to the back room, as shown in table 7.

| SKU | Category | Product Name | Size | Color | Average days to be sold |
|----------------|-----------|--------------|------|-------|-------------------------|
| 77940130040224 | boots | LILLA | 39 | 22 | 162,00 |
| 33430930040325 | cardigan | MALESIA | NULL | 32 | 153,71 |
| 60130330040102 | cloak | TOPAZIO | 19 | 10 | 151,83 |
| 63630930040165 | shirt-top | AMARENA | NULL | 16 | 145,88 |
| 57960533060023 | boots | CARDATO | 36 | 2 | 143,92 |
| 15860233060032 | furtrim | BUONA | NULL | 3 | 143,00 |
| 23430330040794 | cardigan | MIA | NULL | 79 | 142,92 |

Table 7. Slow moving SKUs

Another interesting report is to generate a daily list with the lasts days' high sellers and/or fast movers. The high sellers is an information that could be derived from the inventory counts without using RFID data, but fast movers that could enrich the replenishment reports could only be derived from the RFID data. This list would be useful to depict the garments' remaining pieces in the store area and the remaining pieces in the store floor. It's important for the manager to be able to detect:

- The garments that are not present on the sales floor but are available in the backroom (red color in table 8).
- The garments that should be refilled, as there are only few pieces remaining on the store floor (orange colors in table 8).
- The garments that are not available in the whole store area (grey colors in table 8).

The manager could use this type of information to order from the other stores these garments, and/or to bring on the store floor other garments that have the same characteristics (size, color, category etc.) with the unavailable garments (grey colors).

| SKU | Category | Product Name | Size | Color | Times purchased | Remaining pieces | Store floor pieces |
|----------------|--------------|--------------|------|-------|-----------------|------------------|--------------------|
| 31161137020037 | shirt | MOGOL | 48 | 3 | 10 | 1 | 0 |
| 65140643050601 | bag | KUBO | 36 | 60 | 3 | 11 | 1 |
| 80140440050042 | cloak | 1NADAR | 38 | 4 | 3 | 2 | 1 |
| 60160139060034 | cloak | FARNESE | 42 | 3 | 2 | 3 | 1 |
| 89440240090423 | jerseytop | XMULTIAR | NULL | 42 | 3 | 0 | 0 |
| 10861423060012 | coat | SORTE | 38 | 1 | 2 | 0 | 0 |
| 61340730040743 | longtrousers | RIALTO | 21 | 74 | 2 | 0 | 0 |
| 30860526020015 | coat | LUCERNA | 44 | 1 | 2 | 0 | 0 |

Table 8. SKU daily replenishment list – based on the high sellers and fast movers on 26/08/2014

4 Supporting Decision Making

This section discusses how the knowledge extracted from the above approach could be utilized to support managerial and replenishment reports in fashion and apparel retail stores.

4.1 Managerial Reports

Specifically, the reports concerning the transitions graph (figures 4, 5) could be utilized by the manager not only to identify readers' inconsistencies, but also to have a better visibility of the garments transitions in the store. Also, this kind of information per SKU or even per product category level could be used to predict garments' movements.

Furthermore, reports such as the mean times between the locations and the product transitions per product category (figure 6), or per color, size and SKU, could help managers to have a more detailed visibility of the garments. By categorizing the garment categories (figures 7, 9), or even the sizes and the colors, into fast, medium, slow, and no moving, according to their average time from store floor to check out, and by combining this report with the top selling products (or sizes, or color), the manager could decide the space allotted for each garment category, the number of items of each color, or size should be in the store floor. Also, by combining the above information with the top selling and fast moving SKUs he could identify the SKUs that should always have presence in the sales floor, as these satisfy the shoppers' preferences (tables 3, 4). Last but not least, the reports about the "no moving" SKUs, that are presented at the sales floor, but are for a long period unsold (table 6), could be utilized in cases that there is not enough space on the sales floor and there should be a better allocation of the available sales area. In addition, in these cases if we don't have "no moving" SKUs, we could move the "slow moving" SKUs to the back room (table 7). The above information could be also used by the manager to promote these garments ("garment of the day"), or even to design dynamic pricing strategies to get rid of them. This kind of information could also be used to support replenishment reports.

4.2 Replenishment Reports

Respectively, the SKUs found to be "no moving" or "slow moving" for a long period, could be also be reported in the replenishment report in order to transfer them back to the backroom. Thus, we may gain some space on the sales floor for those garments that are kept in the backroom without having a presence on the sales floor (table 5). In addition, the high selling and fast moving garments which are not present on the sales floor but are available at the backroom, the high selling and fast moving

garments that should be refilled, as there are only few pieces remaining on the store floor (table 8), could also be used to support daily replenishment reports. Last but not least, to avoid out-of-shelf situations, a report with the high selling and fast moving garments that is not available in the whole store area could help the manager either to order them from other stores, and/or to bring garments that have the same characteristics (size, color, category etc.) on the store floor.

5 Conclusions & Discussion

This research looks into item level RFID data captured in retail stores and introduces a new way of utilizing these datasets. We show that RFID technology is more than a tool for obtaining inventory counts in the locations of the retail stores. We develop an approach for RFID-enabled visualization of inventory/ product movements/ flows in the store. We are no longer interested only on the quantities of products positioned on shelves, passing through backroom entrance etc. We reveal the products movements' behavior in order to support decisions ranging from shelves space allocation, dynamic pricing programs for slow-moving fresh products, to product assortment. In the context of SERAMIS EU project, we put this RFID data analytics module in practice with real data provided by a fashion retailer.

We intend to obtain and use more RFID data to support a bigger range of business decisions. For example, we can detect the patterns and the affinities of the garments shoppers try on from the fitting rooms' RFID data. Customers frequently enter a garment store having in their mind (a) specific shopping goal(s), but often they don't purchase the garments they visit the store for. Perhaps the garments didn't fit them, or they didn't find the size they looked for etc. For these reasons, it could be meaningful to compare the garments' patterns and sales affinities derived from the point-of-sales (POS) data, with those derived from the fitting room data, as this kind of data reflects the initial and maybe the real shopping goals of the consumers. Garments patterns derived from the fitting rooms may be different of those derived from the check-out; these could convey signals to the managers that there are selling gaps. Use of RFID technology at the POS level can be used to generate demand trends and build a probabilistic demand pattern. This application is invaluable for retail apparel industry with high levels demand uncertainty (Zhu et al. 2012). Moreover, the fitting room RFID data can reveal those SKUs that are often detected in the fitting room, but they don't reach the check out. Having this kind of information, we can design promotions and even dynamic pricing strategies. In general, the item level RFID data from a real time location system (RTLS) all over the store could enlighten the value and the decisions could be extracted from the system, as it could offer a greater traceability and visibility of the garments at any time (Tijun et. al, 2015; Zhu et al., 2012). Indicatively, some of the advantages are that the manager could obtain a fully inventory visibility, also misplaced garments could be detected etc. Last but not least, by combining the RFID data with other data sources, such as weather data, we could enrich replenishment reports; not accidentally umbrellas are the fast sellers in figure 9.

Acknowledgements

The authors acknowledge SERAMIS project partners (<http://seramis-project.eu>) for providing the necessary data for this study.

REFERENCES

- Bertino, E., Bernstein, P., Agrawal, D., and Davidson, S. (2011). 'Challenges and Opportunities with Big Data', A community white paper developed by leading researchers across the United States.
- Dimakopoulou, A. G., Pramataris, K. C., and Tsekrekos, A. E. (2014). Applying real options to IT investment evaluation: The case of radio frequency identification (RFID) technology in the supply chain, *International Journal of Production Economics*, 156:191–207.

- Hardgrave, B.C., Aloysius, J.A., and Goyal, S. (2013). RFID-Enabled Visibility and Retail Inventory Record Inaccuracy: Experiments in the Field. *Production and Operations Management*, 22 (4): 843–856.
- Hardgrave, B. C., Goyal, S. and John A. (2011). Improving inventory management in the retail store: The effectiveness of RFID tagging across product categories. *Operations Management Research*, 4(1): 6-13.
- Hevner, A. R., March, S. T., Park, J., and Ram, S. (2004). Design science in information systems research. *MIS Quartely*, 28(1): 75-105.
- Keller, T., Thiesse, F., Fleisch, E. (2014). Classification Models for RFID-Based Real-Time Detection of Process Events in the Supply Chain: An Empirical Study. *ACM Transactions on Management Information Systems* 5(4).
- Phan, D. D., and Vogel, D. R. (2010). A model of customer relationship management and business intelligence systems for catalogue and online retailers. *Information & Management*, 47(2): 69-77.
- Provost, F., and Fawcett, T. 2013. Data Science and its Relationship to Big Data and Data-Driven Decision Making *Big Data*, 1(1): 51-59.
- Sellitto C., Burgess S., and Hawking P. (2007). Information quality attributes associated with RFID-derived benefits in the retail supply chain, *International Journal of Retail & Distribution Management*, 35(1): 69-87.
- Sharda, R., Asamoah, D. A., and Ponna, N. (2013). Business analytics: Research and teaching perspectives. In V. Luzar-Stiffler, & I. Jarec (Eds.), *Proceedings of the ITI 2013 35th International Conference on Information Technology Interfaces: 19-27. Cavtat / Dubrovnik, Croatia: IEEE.*
- Thiesse, F., and Buckel T. (2015). A comparison of RFID-based shelf replenishment policies in retail stores under suboptimal read rates. *International Journal of Production Economics*, 159 (1): 126-136.
- Tijun F., Feng T., Sheng D., and Shuxia L. (2015). Impact of RFID technology on supply chain decisions with inventory inaccuracies, *International Journal of Production Economics*, 159: 117-125,
- Unger R., and Sain J. (2015). Kurt Salmon RFID in Retail Study. A retail white paper published in *RFID Journal*. Accessed at <http://www.rfidjournal.com/retail/whitepapers>.
- Zhu, X., Mukhopadhyay, S.K. and Kurata, H. (2012). A review of RFID technology and its managerial applications in different industries. *Journal of Engineering and Technology Management*, 29(1): 152-167.

THE SHAPE OF DIGITAL TRANSFORMATION: A SYSTEMATIC LITERATURE REVIEW

Complete Research

Henriette, Emily, Télécom Ecole de Management, France, emily.henriette@telecom-em.eu

Feki, Mondher, Télécom Ecole de Management, France, mondher.feki@telecom-em.eu

Boughzala, Imed, Télécom Ecole de Management, France, imed.boughzala@telecom-em.eu

Abstract

Digital market has never been so unstable due to more and more demanding users and new disruptive competitors. CEOs from most of industries investigate digitalization opportunities. Through a Systematic Literature Review, we found that digital transformation is more than just a technological shift. According to this study, these transformations have had an impact on the business models, the operational processes and the end-users experience. Considering the richness of this topic, we had proposed a research agenda of digital transformation in a managerial perspective.

Keywords: Digital transformation, Business models, Operational processes, User experience.

1 Introduction

In the past few years, industries are facing technological shifts. Market volatility has resulted in a need for a better response to demand. In a perspective of enabling business agility and changing the way people work to optimize business performance, companies have undertaken digital transformation. Some of the most important innovations are essentially based on internet and cloud technologies; also called digital technologies.

Digital transformation, also known as digitalization, refers to a business model driven by “the changes associated with the application of digital technology in all aspects of human society” (Stolterman and Fors, 2004, p. 689). It is usually implemented through digitization, i.e. the “ability to turn existing products or services into digital variants, and thus offer advantages over tangible product” (Gassmann et al., 2014).

As part of our research, we found that most of existing papers regarding digitalization dealt with technological innovations (e.g. mobile technologies, analytics solutions, etc.), while this subject actually covers a wider potential of scope. We believe that digital transformation should also be studied from a different angle. Indeed, both observations and existing studies from professional papers (MIT - Cap Gemini, 2013; IBM Institute for Business Value, 2012) expose that digital transformation affects every aspect of an organization. The four aspects we will be focusing on are digital capabilities, business models, operational processes and user (internal and external IT consumer) experience.

In this paper, we intend to provide a research agenda on digital transformation with new perspectives, based on systematic literature review method. We will discuss the following questions:

- What are the digital capabilities impacted by the digital transformation?
- How digitalization transforms business models, operational processes and user experience?

The remainder of this paper is organized as follows. Section 2 describes the research methodology. Section 3 presents the results. Finally, section 4 discusses the key findings and provides future research directions.

2 Methodology

The aim of this study is to explore the shape of digital transformation drawing on a literature review. For it, we used a systematic literature review, following Kitchenham (2007) and Okoli and Schabram (2010) protocol. It is a rigorous approach to select, analyze and assess papers. Applied in a given domain, it allows identifying trends and gaps in research.

The systematic literature review follows these following 6 steps:

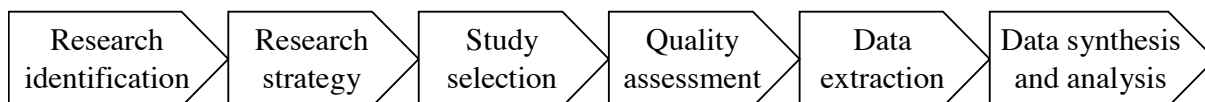


Figure 1. Systematic literature review method.

2.1 Research identification

The idea here is to examine and evaluate research on digital transformation. For that, we investigated the above research questions.

2.2 Research strategy

Our search strategy consists first in deriving major terms related to the research questions, and then identifying alternative spelling and synonyms for these terms by leading a pilot test. We used the Boo-

lean operators (OR; AND) for connecting the founded terms. This resulted in the following used strings for automated search:

("digital transformation" OR "digitalization") AND ("user experience" OR "operational process" OR "business model").

The search of articles was conducted regardless of time limitation of publications by using Scopus database. This bibliographic database holds more than 21,000 peer-reviewed journals, over 1,200 “open access” journals, more than 600 trade publications, 350 book series. The search of articles has begun on May 16th, 2015.

2.3 Study selection

In this step, we defined selection criteria to determine which studies are included or excluded. Studies that met the following criteria were included:

- The paper should be written in English
- The paper should be published in a scientific journal
- The paper approaches digital transformation

The articles which they weren't accessible stated as excluded, as well as, master and doctoral theses, proceedings or conference articles, working papers and textbooks. This choice of journal articles falls in line with Ngai and Wat (2002, p.416), who believe that “academics and practitioners alike use journals most often for acquiring information and disseminating new findings and represent the highest level of research”.

The final list of considered publications included 202 articles. Both authors carried out the selected study process independently. Each reviewer performed the screening of the results based on title and abstract for each publication that was considered according to the inclusion and exclusion criteria. Then, a comparison of screening results is realized, in case of difference, verification is jointly made to reach a consensus. At the end of this process, 153 articles were excluded and 49 articles were kept for the quality assessment step.

2.4 Quality assessment

In this step, the quality criteria are defined to evaluate the rigor and credibility of the selected articles. The evaluation requires the complete review of the paper. Based on the works of Nguyen-Duc et al. (2015), Hauge et al. (2010), and Dyba and Dingsoyr (2008), we defined the following quality stated criteria as questions:

- Is there an adequate description of the context in which the research was carried out?
- Is there a clear statement of research aims?
- Does the paper describe an explicit research question?
- Is the research design appropriate to address the research aims?
- Is the literature review adequate?
- Is the collected data in a way of addressed research issue?
- Is the data analysis sufficiently rigorous?
- Is there a clear statement of findings?
- Is the study valuable for research or practice?
- Does the paper discuss limitations or validity?

Each question has four possible options: (0) issue is not mentioned at all, (1) little mentioned, (2) adequately addressed and (3) completely addressed (Nguyen-Duc et al., 2015). Hence, we used a four points Likert scale for collecting answers. Articles with an average quality score lower than 1, were removed. At the end of this process 13 articles were qualified to be analyzed for the data extraction step.

2.5 Data extraction

In this step, we extracted data from the qualified articles.

2.6 Data synthesis and analysis

At the end, some results came out of the extracted data. The data synthesis includes a descriptive analysis to provide a background about the included articles and an analysis of their findings in order to underline the future directions of research.

Figure 2 presents the literature search, selection and assessment process.

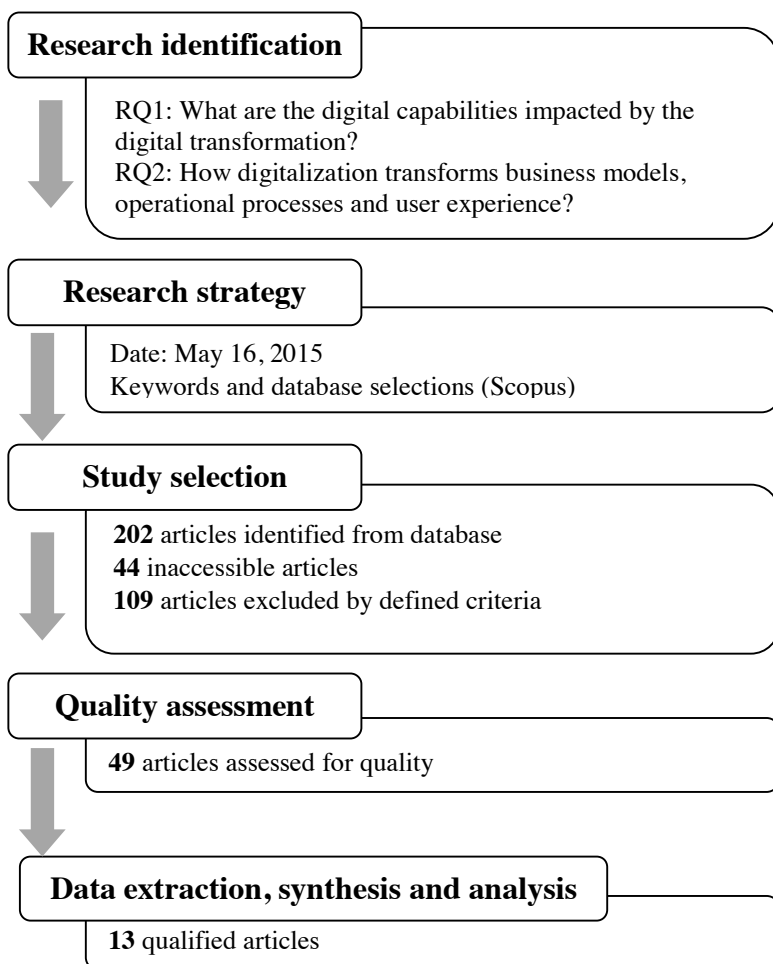


Figure 2. Systematic literature review process.

3 Results and Analysis

This section presents an overview of 13 selected articles and a classification by digital capabilities, business model, operational process, and user experience.

3.1 Distribution of articles by year of publication

Table 1 presents the distribution of articles by the year of publication. The first article was published in 2006. From 2011 to 2012, the amount of articles doubled each year to finally stabilize to 3 articles per year in 2013 and 2014. As the subject is contemporary and at leading-edge, publications are steady over time.

| Year | Number of articles (%) | References |
|-------|------------------------|--|
| 2006 | 1 (8%) | (Zhu et al., 2006) |
| 2011 | 2 (15%) | (Kohli et al., 2011); (Rogers et al., 2011) |
| 2012 | 4 (30%) | (Berman, 2012); (Gastaldi et al., 2012); (Liu, 2012); (Pinzaru et al., 2012) |
| 2013 | 3 (23%) | (Barland, 2013); (Belk, 2013); (Medina et al., 2013) |
| 2014 | 3 (23%) | (Pardo et al., 2014); (Øiestad et al., 2014); (Rothmann et al., 2014) |
| TOTAL | 13 (100%) | |

Table 1. Distribution of articles by year of publication.

3.2 Distribution of articles by journal

Figure 3 represents the distribution of articles by journal subject area. We can notice that digitalization covers a lot of areas such as social sciences, information system and management.

5 articles are published in a high ranked journal:

- MIS Quarterly Executive (Kohli et al., 2011)
- European Journal of Information Systems (Zhu et al., 2006)
- Technological Forecasting & Social Change (Øiestad et al., 2014), (Rothmann et al., 2014)
- Journal of Consumer Research (Belk, 2013)

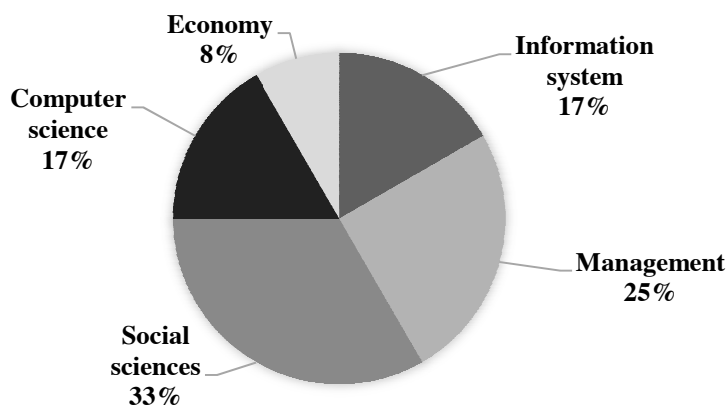


Figure 3. Articles distribution by journal subject area.

3.3 Distribution of articles by methodology

The distribution of articles by research methodology (Palvia et al., 2006) is shown in Table 2. Most of the articles (6 – 60%) use “case study” methodology due to the contemporaneity of the subject and because most of the articles have for main subject digital business models (of which case study is an appropriate method). 20 % of articles adopt a frameworks and conceptual model development (2 articles). The rest of articles use quantitative and qualitative research (with 1 article each).

We also noticed that most of the time, research methodology was poorly developed and research limits were not expressed. Moreover, we couldn't find any rigorous literature review dealing with the concept of digitalization.

| Methodology | Number of articles (%) | References |
|---------------------------------|------------------------|--|
| Case study | 6 (60%) | (Liu, 2012); (Rothmann et al., 2014); (Barland, 2013); (Øiestad et al., 2014); (Kohli et al., 2011); (Gastaldi et al., 2012) |
| Frameworks and conceptual model | 2 (20%) | (Zhu et al. 2006); (Pînzaru et al., 2012) |
| Quantitative research | 1 (10%) | (Pardo et al., 2014) |
| Qualitative research | 1 (10%) | (Medina et al., 2013) |
| TOTAL | 10 (100%) | |

Table 2. Distribution of articles by methodology.

3.4 Distribution of articles by application area

Table 3 presents the studies which were applied on industries. A large part of articles deals with non-material industries. We can suppose that researchers have more hindsight on both internet-based and digitization business models which are the pioneers of digital era thus bring more perspectives for a case study.

| Industry | Number of articles (%) | References |
|-------------------|------------------------|--|
| Book / publishing | 3 (30%) | (Liu, 2012); (Øiestad et al., 2014); (Rothmann et al., 2014) |
| Media | 3 (30%) | (Barland, 2013); (Medina et al., 2013); (Pardo et al., 2014) |
| Music | 2 (20%) | (Pînzaru et al., 2012); (Rogers et al., 2011) |
| Healthcare | 1 (10%) | (Gastaldi et al., 2012) |
| Oil/Gas | 1 (10%) | (Kohli et al., 2011) |
| TOTAL | 10 (100%) | |

Table 3. Distribution of articles by application area.

3.5 Classification of articles by digital capabilities

A firm's IT capabilities represent "the application of physical or intangible IT resources such as technology, knowledge, practices, relationships, management skills, business process understanding and human resources to further organizational goals" (Sandberg et al., 2014). We made the assumption that a digital capability is a kind of IT capability. In digital context, the main difficulty is to identify the boundaries of digital capabilities. None of the author tried to neither identify these boundaries nor define what a digital capability is.

However, they mentioned different capabilities presented on table 4. Among digital capabilities, we can identify technological assets which purposes are to optimize and transform business activities such as analytics and mobility.

Digitalization can be implemented through the transformation of physical asset to digital asset (digitization). For example, the book industry has lived a digital shift with the dematerialization from physical books to e-books (Liu, 2012). Internet and social network enable the access of a new field of potential customers and strengthen ties with existing customers (Berman, 2012)

Digital transformation also has an organizational impact on human resources. Job roles evolve in line with the transformation of activities. Decision makers must have to take into account the evolution of knowledge and skills (Kohli et al., 2011; Liu, 2012).

| Digital capability | Number of articles (%) | References |
|----------------------------------|------------------------|--|
| Digitization / dematerialization | 5 (18%) | (Liu, 2012); (Rothmann et al. 2014); (Øiestad et al., 2014); (Belk, 2013); (Gastaldi et al., 2012) |
| Internet technologies | 10 (37%) | (Liu, 2012); (Rothmann et al., 2014); (Kohli et al., 2011); (Berman, 2012); (Øiestad et al., 2014); (Zhu et al., 2006); (Pardo et al., 2014); (Rogers et al., 2011); (Pınzaru et al., 2012); (Medina et al., 2013) |
| Analytics | 3 (10%) | (Kohli et al., 2011); (Berman, 2012); (Gastaldi et al., 2012) |
| Mobility | 5 (18%) | (Berman, 2012); (Liu, 2012); (Pardo et al. 2014); (Pınzaru et al., 2012); (Medina et al., 2013) |
| Social Network | 2 (7%) | (Berman, 2012); (Rogers et al., 2011) |
| Knowledge and skills | 3 (10%) | (Kohli et al., 2011); (Liu, 2012); (Belk, 2013) |
| TOTAL | 28 (100%) | |

Table 4. Distribution of articles by digital capability.

3.6 Classification of articles by business model

In this paper's context, a business model is "a description of a company's intention to create and capture value by linking new technological environments to business strategies" (Liu, 2012). We found that digitalization can be considered either as a business model or as a reshaping of existing business model, taking into account digital capabilities (Rothmann et al., 2014).

Along with the technological shift, convergence of social media and mobile technologies is changing the way of conducting business. Opportunities from new capabilities (for example, dematerialization) are a door opener to an extended market and result to an adjustment of business focus. We also noticed that digitalization is often presented as an inescapable evolution as market imperatives present a high risk of not doing the technological shift, especially for publishing and music companies.

Table 5 presents the distribution of articles by impact on business model. We identified 3 impacts on business models: Extend market (4 articles -36%), focusing on customer value propositions (3 articles – 28%) and reshaping existing business model due to market imperatives (4 articles – 36%).

| Business model | Number of articles (%) | References |
|---|------------------------|---|
| Extend market | 4 (36%) | (Liu, 2012); (Øiestad et al., 2014); (Pînzaru et al., 2012); (Medina et al., 2013) |
| Focusing on customer value propositions | 3 (28%) | (Berman, 2012); (Rogers et al., 2011); (Pînzaru et al., 2012) |
| Reshaping existing business model due to market imperatives | 4 (36%) | (Rothmann et al., 2014); (Kohli et al., 2011); (Rogers et al., 2011); (Medina et al., 2013) |
| TOTAL | 11 (100%) | |

Table 5. Distribution of articles by impact of digitalization on business model.

3.7 Classification of articles by user experience

Table 6 presents the distribution of articles by impact of digitalization on user experience. The majority of articles based on user maturity (6 articles – 46%), followed by interaction (3 articles – 15%) and collaboration (2 articles – 18%).

Our review showed that the user is in the heart of digital transformations. Customers are more demanding and expect companies to listen, understand and be flexible about the evolution of their needs. In companies, users, by using new technologies in private context, expect to use the same technologies at work. Especially for new generations, also known as “digital natives” (Pardo et al., 2014), which are born surrounded by technologies.

Collaboration tools expansion and commoditization of social networks changed the way of working by tightening interactions between users and their ecosystem.

| User experience | Number of articles (%) | References |
|---------------------------------|------------------------|---|
| Digital natives & user maturity | 6 (54%) | (Rothmann et al. 2014); (Berman, 2012); (Øiestad et al., 2014); (Pardo et al., 2014); (Rogers et al., 2011); (Pînzaru et al., 2012) |
| Interaction | 3 (28%) | (Berman, 2012); (Belk, 2013); (Pînzaru et al., 2012) |
| Collaboration | 2 (18%) | (Berman, 2012); (Belk, 2013) |
| TOTAL | 11 (100%) | |

Table 6. Distribution of articles by impact of digitalization on user experience.

3.8 Classification of articles by operational process

Undertaking digital transformation can impact the whole company supply chain. Table 7 presents the distribution of articles by operational process impacted by digitalization.

Our review showed two kinds of impact in operational processes:

- Transformation of processes induced by implementation of new technologies: As we explained above, digitalization transforms the way to work, both in term of user experience and business model. For example, implementing analytic tools to make predictive analysis on consumption trends will impact the way of doing marketing, and so, the marketing process (Berman, 2012).
- Digitalization of a chosen process: Companies can decide to undertake the digitalization of a specific process which implies to make investments in order to modernize a full process. For example the digitalization of knowledge management implies to invest on new technologies (such as collaboration or analytic tools), to conduct change on usages and assign specific resources on knowledge management (Kohli et al., 2011).

| Operational process | Number of articles (%) | References |
|-----------------------|------------------------|---|
| Supplier relationship | 1 (10%) | (Kohli et al., 2011) |
| Customer relationship | 1 (10%) | (Kohli et al., 2011) |
| Knowledge management | 1 (10%) | (Kohli et al., 2011) |
| Marketing | 1 (10%) | (Berman, 2012) |
| Delivery | 3 (30%) | (Kohli et al., 2011); (Berman, 2012) |
| Sales / Engagement | 1 (10%) | (Berman, 2012) |
| Knowledge management | 2 (20%) | (Kohli et al., 2011); (Gastaldi et al., 2012) |
| TOTAL | 10 (100%) | |

Table 7. Distribution of articles by operational process impacted by digitalization.

4 Discussion and future research directions

We have seen that digital transformation is often described as a new business model or as a reshaping of existing business models. It is driven by several factors:

- Companies are vulnerable to these new technologies: the past few years of technological shifts have broken down market barriers for new disruptive competitors (e.g. Netflix, Uber, etc). Industries such as publishing, media or music had to undertake deep transformations, especially by digitizing assets.
- Opportunities to extend the market: we saw that new capabilities (e.g. internet and mobile technologies) are a door opener to an extended market. They require a company to adjust their business model accordingly.

- A high and fluctuant expectation from users: with a deep knowledge on new technologies users wish to use personalized and cutting-edge technologies.

A digital transformation project involves implementing digital capabilities to support business model transformations. It impacts the whole organization, especially operational processes, resources, internal and external users. This is a major change in ones habits and ways of working, which is based on collaboration and intensive interactions.

Table 8 summarizes findings of this study, responding to our following research questions: What are the digital capabilities impacted by the digital transformation? How digitalization transforms business models, operational processes and user experience?

| Research questions | | Findings |
|---|-----------------------|--|
| What are the digital capabilities impacted by the digital transformation? | | Digitization / dematerialization; Internet technologies; Analytics; Mobility; Social Network; Knowledge and skills |
| How digitalization transforms business models, operational processes and user experience? | Business models | Extend market; Focus on customer value propositions; Reshaping existing business model due to market imperatives |
| | Operational processes | Supplier relationship; Customer relationship; Knowledge management; Marketing; Delivery; Sales / Engagement |
| | User experience | Digital natives & user maturity; Collaboration; Interactions |

Table 8. Findings summary.

Nonetheless, our research is limited by several factors such as the chosen keywords. We would have, for example, added afterwards other keywords such as “consumer experience” or “business process”. We can add others databases such as “science direct” or “business source complete” in order to enrich our results. Regardless of the prior, this first investigation also highlights some future research directions.

First, we believe that the priority is to settle a rigorous theoretical frame on what digitalization is and what digital capabilities are.

Second, we identified a lack of research regarding the realization of digital transformation projects. The questions that arise are: How to manage a digital transformation? ; How to identify and manage the costs of this transformation?

Third, it would be interesting to make a digitalization maturity assessment tool to identify opportunities and make a benchmark of organizations interested in leading a digital transformation. Then, a guide for digital transformation could help industries initiate a change and frame their project.

Fourth, as we have seen above, digitalization projects have a strong impact on this whole organization. The authors mentioned some of them; however, a research focused on digitalization impacts would also be interesting; especially regarding its impacts on IT. This has an even larger impact on big companies. Indeed, blue-chip companies have a significant and aging infrastructure (system of record), which might have to evolve to adjust to these digital technologies. It also raises issues regarding IT management and governance: who will manage IT infrastructure, how to avoid shadow IT and how IT jobs will evolve?

Finally, our three research questions had resulted in that digitalization impacts business models, operational processes and the user experience. Research on these three areas should be deepened. To lead proposed research, we suggest the use of different approaches (e.g. design science or quantitative approach) and expand application area (blue-chip companies, other industries such as retail or governments).

References

- Barland, J. (2013). "Innovation of New Revenue Streams in Digital Media." *Nordicom Review* 34, 99-112.
- Belk, R. (2013). "Extended Self in a Digital World." *Journal of Consumer Research* 40 (3), 477-500.
- Berman, S. (2012). "Digital transformation: opportunities to create new business models." *Strategy & Leadership* 40 (2), 16-24.
- Dyba, T. and T. Dingsoyr (2008). "Empirical Studies of Agile Software Development: A Systematic Review". *Information and Software Technology* 50 (9-10), 833-859.
- Gassmann, O., Frankenberger K. and M. Csik (2014). *The St. Gallen Business Model Navigator*. URL: http://www.im.ethz.ch/education/HS13/MIS13/Business_Model_Navigator.pdf
- Gastaldi, L. and M. Corso (2012). "Smart Healthcare Digitalization: Using ICT to Effectively Balance Exploration and Exploitation Within Hospitals." *International Journal of Engineering Business Management* (4), 1-13.
- Hauge, O., Ayala C. and R. Conradi (2010). "Adoption of Open Source Software in Software-intensive Organizations – A Systematic Literature Review." *Information and Software Technology* 52 (11), 1133–1154.
- IBM Institute for Business Value (2012). *CMOs and CIOs Acquaintances or allies?*. URL: https://www.ibm.com/smarterplanet/global/files/se_sv_se_intelligence_CMOs_and_CIOs.pdf
- Kitchenham, B. (2007). *Guidelines for Performing Systematic Literature Reviews in Software*. Technical Report EBSE-2007-01, UK, Keele University and University of Durham.
- Kohli, R. and S. Johnson (2011). "Digital Transformation in Latecomer Industries: CIO and CEO Leadership Lessons from Encana Oil & Gas (USA) Inc." *MIS Quarterly Executive* 10 (4), 141–156.
- Liu, D. (2012). "Competitive Business Model in Audio-book Industry: A Case of China." *Journal of software* 7 (1), 33-40.
- Medina, M. and B. Prario (2013). "The transformation of audiovisual media companies: The cases of Mediaset (Italy) and Antena 3 (Spain)." *Studies in Communication Sciences* 13 (2), 166–173.
- MIT - Cap Gemini (2013). *Digital transformation: a roadmap for billion-dollar organizations*. URL: https://www.capgemini.com/resource-file-access/resource/pdf/Digital_Transformation_A_Road-Map_for_Billion-Dollar_Organizations.pdf
- Ngai, E. W. T. and F. K. T. Wat (2002). "A literature review and classification of electronic commerce research." *Information & Management* 39 (5), 415–429.
- Nguyen-Duc, A., Cruzes, D. S. and R. Conradi (2015). "The impact of Global Dispersion on Coordination, Team Performance and Software Quality: A Systematic Literature Review." *Information and Software Technology* (57), 277-294.
- Øiestad, S. and M. Bu (2014). "Digitisation of publishing: Exploration based on existing business models." *Technological Forecasting & Social Change* 83, 54-65.
- Okoli, C. and K. Schabram (2010). "A Guide to Conducting a Systematic Literature Review of Information Systems." *Sprouts: Working Papers on Information Systems* 10 (26). <http://sprouts.aisnet.org/10-26>
- Palvia, P., V. Midha and P. Pinjani (2006). "Research Models in Information Systems." *Communications of the Association for Information Systems* 17, 1042-1063.
- Pardo, A. and C. Etay (2014). "Movies and screens: the Spanish audience's choice." *Communication & Society* 27 (4), 131-145.
- Pinzaru, F. and A. Mitan (2012). "Social media and marketing of the "popcorn" music wave: the success of romanian commercial musicians analysed through their perceived image on facebook and youtube." *Economics & Sociology* 5 (2a), 125-138.
- Rogers, J. and S. Sparvierio (2011). "Same tune, different words: The creative destruction of the music." *Observatorio Journal* 5 (4), 1-30.

- Rothmann, W. and J. Koch (2014). “Creativity in strategic lock-ins: The newspaper industry and the digital revolution”. *Technological Forecasting & Social Change* 83, 66–83.
- Sandberg, J., Mathiassen, L. and N. Napier (2014). “Digital Options Theory for IT Capability Investment.” *Journal of the Association for Information Systems* 15 (7), 422–453.
- Stolterman, E. and A. C. Fors (2004). “Information Technology and the Good Life.” In: *Information Systems Research: Relevant Theory and Informed Practice*. Ed. Kaplan, B. et al., London: Kluwer Academic Publishers.
- Zhu , K., Dong, S., Xin Xu, S. and K. Kraem (2006). “Innovation diffusion in global contexts: determinants of post-adoption digital transformation of European companies.” *European Journal of Information Systems* 15, 601–616.

MOBILE APPS FOR OMNICHANNEL RETAILING: REVEALING THE EMERGING SHOWROOMING PHENOMENON

Complete Research

Lazaris Chris, ELTRUN - The E-Business Research Center, Athens University of Economics & Business, School of Business, Department of Management Science & Technology, Greece, lazaris@aueb.gr

Vrechopoulos Adam, ELTRUN - The E-Business Research Center, Athens University of Economics & Business, School of Business, Department of Management Science & Technology, Greece, avrehop@aueb.gr

Doukidis Georgios, ELTRUN - The E-Business Research Center, Athens University of Economics & Business, School of Business, Department of Management Science & Technology, Greece, gjd@aueb.gr

Fraidaki Katerina, ELTRUN - The E-Business Research Center, Athens University of Economics & Business, School of Business, Department of Management Science & Technology, Greece, fraidaki@aueb.gr

Abstract

The transformation of the smartphone into a key integrating factor of the online & offline retailing environment has led to the development of mobile applications that shape the omniretailing landscape. The present study provides evidence of the mobile retailing apps frequency of use inside physical stores and explores mobile retailing app assisted shoppers' preferences of in-store omniretailing practices & technologies. Results reveal that price comparison that could lead to showrooming is of utmost important for consumers. In parallel, consumers that attach great importance to such practice significantly differ from the rest, in terms of the importance they attach to salespeople & omnichannel integration criteria, in order to purchase offline. In contrast, there weren't found statistically significant differences in terms of the importance they attach to online & offline store atmosphere. Nevertheless, the importance attached to online store atmosphere is high among mobile retailing app assisted shoppers. Drawing on these results, the study provides feedback to retail entrepreneurs regarding the optimal design and features of the future physical retail store.

Keywords: Omnichannel Retailing, Mobile Apps, Consumer Behaviour, Showrooming, Store Atmosphere.

1 Introduction

It is evident that smartphones have become an important part of everyday life. They could be described as life companions, since users seem to integrate them into their daily activities¹. This phenomenon can be attributed to several factors. First of all, having internet access at all times provides value added services that enhance the users' physical activities. In addition to this, users are able to

¹ <http://www.emarketer.com/Article/Smartphones-In-Store-Shopping-Companions/1010800>

take advantage of hardware-assisted internet features that smartphones provide them, such as sensors, location based services and cameras. Finally, the combination of always-on internet access and hardware supported features is complemented with new user interfaces that outperform the conventional web environment experience provided by modern web browsers: the mobile applications (apps) GUI. Mobile apps seem to be the driving force of smartphones, creating ecosystems that engage users and influence their behaviour.

From a retailing perspective, smartphones play an important role integrating retailing channels, blending online with offline, since consumers inside physical stores are at the same time mobile-assisted online shoppers. Multichannel retailing transforms into a complex, diversified form of retailing, recently characterized as “omnichannel retailing” or “omniretailing”. On the one hand, reports² show that physical retail stores will continue to be consumers’ preferred point of purchase and that online sales will only account for a small portion of total sales. While there are several reasons for that forecast, an obvious one is the clear superiority of the physical environment in comparison to the online one: it attracts more physical senses, with tactile being the most important one, according to lab experiments (Spence & Gallace, 2011).

Conversely, e-commerce provides unique benefits to shoppers that are absent from the physical store. Online features such as instant price comparison, fast checkouts, recommender systems and product reviews accessibility are quite popular in e-tailing. In early m-commerce era, such practices were impossible or difficult to perform within the physical store since mobile phones were not smart enough (software and hardware-wise) for consumers to take fully advantage of them. Whilst smartphones’ hardware specifications continually evolve along with innovative software features in the form of mobile apps, they provide consumers with a convenient access to the online retailing environment, inside the physical one, transforming them into “omnishoppers”.

The purpose of this study is to explore the consumers’ retailing mobile app & in-store internet penetration within the physical store. Furthermore, the study attempts to clarify which omniretailing technologies & practices are most important for mobile app assisted in-store internet users. Finally, consumers’ preferences are analyzed providing feedback for retailing entrepreneurs that are interested in designing the future retail store and enhancing it with omniretailing features.

2 Literature Review, Research Hypotheses and Methodology

At first, internet services were meant to be utilized by the use of internet browsers. Early browsers were optimized for the desktop environment, whereas web content consisted mainly of document files (html) and few multimedia elements (e.g. images, audio). Next, the Web 2.0 era emerged as the internet experience included dynamic web pages, asynchronous network communication and richer content, converting the web into an application environment (Mikkonen & Taivalsaari, 2011). Since the introduction of the mobile as a new internet access device, several attempts were made in order to transfer the internet experience on the move. Early mobile phones featured small, non-touch screens and low hardware specs which resulted in poor internet browsing, which followed the WAP protocol. As the devices became more powerful and their screens improved, both in size and in quality, the internet experience could be offered by html browsers. Still, usability remained an issue, since small keyboards, or even touch pens could not provide efficient and effective human-computer interaction.

It was the advent of the Apple iPhone that brought true revolution to this domain. The success of the iPhone was not only because of its superior hardware (capacitive multi-touch screen, sensors, etc), but also because of the mobile apps ecosystem it introduced, creating the true smartphone. Users were now able to benefit from online services not only by using the internet browser, but also by utilizing

2 <http://www.emarketer.com/Article/Retail-Sales-Worldwide-Will-Top-22-Trillion-This-Year/1011765?ecid=1001>

mobile applications, downloaded from the application store of the platform. Mobile apps soon became a strong alternative to web sites, and the latter also became mobile friendly. Nowadays, online services can be encountered in the following variants in smartphones: Standard web site (desktop version), Web site featuring responsive design (desktop version adapts to smaller screens), Mobile web site (separate from the desktop version), Native mobile app and Hybrid mobile app (web pages wrapped into native mobile app). Overall, the key differences between mobile app and web site implementations are the following:

- Mobile apps run compiled code, written in C & Objective C (Apple iOS), Java (Google Android), .NET (Windows Phone), whereas mobile web sites typically utilize mobile frameworks, running interpreted code (Charland & Leroux, 2011). Therefore, mobile apps perform faster than mobile web sites (Huy & van Thanh, 2012).
- Mobile apps are device-specific and difficult to implement and maintain, mobile web sites, on the other hand, are typically cross-platform and can be instantly updated (Wisniewski, 2011).
- Mobile apps offer deep mobile OS integration (e.g. alerts and notifications), featuring specific APIs that access device hardware (sensors, cameras, gps, etc) directly, whereas web sites have limited hardware API support, although HTML5 seems to gradually adapt to this situation (Charland & Leroux, 2011; Wisniewski, 2011).
- Mobile apps provide superior user interface (Charland & Leroux, 2011), suitable even for one-handed operation, featuring hardware acceleration and customized software buttons and gestures, web sites, conversely, rely on the web browser interface in order to interact with the user.

Based on the previous differences, it can be assumed that for retailing purposes mobile apps seem to be a more suitable choice for consumers in-store, since they can assist users with more natural interaction (e.g. augmented reality, camera-based product recognition), less clicks featuring one-hand operation (retailing optimized interface), employing more hardware functions (e.g. sensors, bluetooth), and with faster response (Wisniewski, 2011; Mikkonen & Taivalsaari, 2011). In addition to this, mobile apps seem to be the most appropriate way to seamlessly integrate online & offline features, due to direct hardware API & OS support (Huy & van Thanh, 2012). In fact, this could be the key point achieving omnichannel state within the physical store.

Omnichannel stems from the latin word *omnis* (meaning: all, everything) and it was first introduced by practitioners in order to differentiate from multichannel. The concept was that consumers utilize retailing channels simultaneously and not just in parallel (Parker & Hand, 2009; Ortis & Casoli, 2009). In academic literature, it was first encountered by Rigby (2011, p.4) who defined omnichannel retailing as “an integrated sales experience that melds the advantages of physical stores with the information-rich experience of online shopping”. Besides, Omniretailing was introduced as “a coordinated multichannel offering that provides a seamless experience when using all of the retailer’s shopping channels” (Levy, et al., 2013, p.67). Recently, Fairchild (2014, p.1) states that “omnichannel commerce involves combining traditional commerce with online commerce by integrating processes in a harmonious and complementary way throughout the organizational and IT chain, and includes external logistics partners in these processes”. Finally, recent omnichannel-specific literature poses specific mention to mobile apps and the mobile channel referring to it as a “disruptive change in the retail environment” (Verhoef et al., 2015). Consequently, omnichannel includes several aspects of retailing ranging from the consumer point to the retailer or even the whole supply chain. In the remaining of this study we attempt to explore the consumers’ perspective regarding the simultaneous use of channels, inside the physical retail store, utilizing mobile apps as the key integrating technology of the online & offline environment.

There are several studies & reports that depict consumers’ omniretailing practices within the physical store. Some refer to them as mobile-assisted shoppers (Luo et al., 2014; Quint et al., 2013) focusing on

consumers' efforts to assist themselves in-store. Other focus on the online practices they use (Wurmser, 2014; Lazaris et al., 2015) and there are also other studies that attempt to explain their behaviour (Agrebi & Jallais, 2015; Lazaris et al., 2014). All studies and reports agree that mobile plays an important role in in-store shopping and that there is a growing percentage of consumers that adopts omniretailing practices (Adobe, 2013). Following these directions, our initial research question is whether consumers utilize not only mobile internet in-store but also mobile apps. Also, if there is a relationship between in-store internet frequency of use and retailing mobile apps frequency of use by shoppers. Related literature shows that enjoyment, behavioural intention to use mobile internet, educational level, subscription of a flat rate and ease of use are correlated with mobile internet usage criteria (Gerpott & Thomas, 2014). Ease of use is attributed to mobile apps (Charland & Leroux, 2011) and therefore we could assume that apps correlate with mobile internet usage criteria. In addition to this, it was found that mobile apps increase internet traffic to the provider's corresponding mobile website, and therefore mobile internet use (Xu et al., 2014). For that reason, if mobile apps were used in-store, it would also lead to increased in-store internet use. Therefore, the following research hypothesis is formulated:

H1: *There are statistically significant differences between shoppers with different levels of retailing mobile app frequency of use, in terms of their in-store Internet frequency of use levels.*

An aftermath of this hypothesis is what retailing-assisting mobile app users want to do with internet in-store. Previous insights suggest that they want to engage in omniretailing practices utilizing e-commerce technologies that they are familiar with from the online channel, seeking for the omnichannel experience. But which online practices & technologies are most important for them? In a recent related business report, several omniretailing practices inside physical store are presented, with price comparison appearing to be the most favourite (Wurmser, 2014, p.11). The report referred to price comparison in-store, in relation to showrooming. Quint et al. (2013) were among the first that presented this topic in a report entitled: "Showrooming and the Rise of the Mobile-Assisted Shopper", where they also enlist other accompanying consumer practices. Showrooming, was only recently defined in academia by Rapp et al. (2015, p.360) as "a practice whereby consumers visit a brick-and-mortar retail store to (1) evaluate products/services firsthand and (2) use mobile technology while in-store to compare products for potential purchase via any number of channels". The study investigated the role of the salesperson towards this behaviour. Similarly, Luo et al. (2014) examined the showrooming intention of mobile-assisted shoppers in a multichannel retailing environment, regarding it as an important phenomenon, with pricing and employee knowledge competency to play an important role in it. At the same time, Willmott (2014) presented several statistical findings and reports that showrooming goes mainstream among mobile shoppers as a common practice.

Nonetheless, price comparison was also a favourite online practice years ago, when Burke (2002) investigated 128 different aspects of the shopping experience online & in-store, conducting a national survey with 2.120 online users. Price comparison online was considered "must have" for 28,1% of respondents and "should have" for 66,9% of them. Although the study incorporated in-store shopping features, price comparison was not included among them at that time probably because neither smartphones nor efficient online price comparison shopping engines existed. Also, price comparison was not included by Mahatanankoon et al. (2005) who explored consumer perception of 44 mobile applications at early days of m-commerce. Apart from price comparison and showrooming, several other research papers offer recommendations about mobile app features. Zhao & Balagué (2015) provided recommendations for branded mobile apps features and categorized them in tool-centric, game-centric, social-centric, m-commerce centric & design centric. Similarly, Magrath and McCormick (2013) presented a product & services design m-marketing design framework depicting several features for mobile fashion retail apps. Based on the previous studies and business reports, we selected 18 online practices and technologies that are compatible with omnichannel retailing, in order to rank and explore the importance that mobile retailing app consumers attach to them in-store. Consequently, based on previous literature, the following research hypotheses are formulated:

H2: *There are statistically significant differences across a series of online practices & technologies applied inside physical stores in terms of the importance that in-store Internet retail mobile app assisted shoppers attach to them.*

In addition to this, we postulate that between retail mobile app and non mobile app assisted in-store internet shoppers significant differences exist. The reason for this is the nature of mobile apps, the additional hardware-assisted features that they support and the overall differentiated smartphone features, as presented above. Therefore, we also propose the subsequent hypothesis:

H3: *There are statistically significant differences between retail mobile app and non mobile app assisted in-store Internet shoppers in terms of the importance they attach to a series of online practices & technologies applied inside physical stores (18 sub-hypotheses in total, as the number of online practices & technologies).*

Next, a subsequent research question emerges: if price comparison and more specifically showrooming intention is high among in-store internet users, what could be done to prevent it? This subject remains an open issue and literature reveals several approaches that could be followed. Chiu et al. (2011) provided 3 effects that have an impact on cross-channel free-riding behaviour, a term similar to showrooming: searching for product information in retail channel and then purchasing it in another one (Chiu et al., 2011, p.1). According to their study, the “push” effect is consumers’ perceived multichannel self-efficacy that positively influences showrooming. The “pull” effect is the attractiveness of competitor’s physical retail store, which also has a positive effect on showrooming. In other words, a consumer may leave a store in order to purchase from another one which has more attractive store atmosphere. Finally, the “mooring” effect, which are lock-in levels within the retailer negatively impact showrooming. That is, factors that make it difficult for the consumer to switch to another retailer (e.g. time consuming or involving complicated procedures). Nevertheless, we should mention that this study only examined free-riding from the online channel to the offline one (research online, purchase offline). Next, Shukla & Babin (2013) discovered that regarding store switching behaviour, hedonic values are more important than utilitarian ones and, therefore, retailers should pay attention to the retail store environment in order to reduce consumer defection. In contrast, Heitz-Spahn (2013) addressed three motives to cross-channel free-riding behaviour: shopping convenience, flexibility and price comparison. Interestingly, they discover that channel aesthetics as components of store atmosphere, although important, do not influence retailer & channel choice and therefore showrooming. They also suggested that utilitarian motives (e.g. pricing) are more important than hedonic ones (e.g. design, ergonomics) towards this issue. They also proclaimed that mobile applications are turning to be a significant research direction towards this area.

At this point it should be noted that store atmosphere notion is applicable both online and offline, with different components and definitions characterizing it throughout the years. Eroglu & Machleit (1993), reported that store atmospherics consist of “all of the physical and non-physical elements of a store that can be controlled in order to enhance (or restrain) the behaviors of its occupants, both customers and employees”. In parallel, Dailey (2004, p.796) stated that a web atmospheric cue is “comparable to a brick-and-mortar atmospheric cue and can be defined as any web interface component within an individual’s perceptual field that stimulates one’s senses”. In fact, atmospherics also extend to the mobile domain in the form of m-atmospherics (Manganari et al., 2007).

Only recently, Pantano & Viassone (2015) considered store atmosphere & channels availability to impact purchase intention. These factors are also found to affect service quality perception, which is also affected by technology and/or salesperson interaction. The study concluded indicating that consumers evaluate all channels simultaneously and therefore retailers should integrate them seamlessly through the use of mobile technologies such as iBeacon, mobile apps and smartphones. In fact, they suggested that multichannel integration is the right step towards avoiding cross-channel free riding behaviour. Also, regarding multichannel integration, Zhang & Oh (2013) exploring customer switching behavior, proposed that retailers should focus on providing innovative cross-channel services in order to retain customers and enhancing service convenience. As far as service is concerned, Monteleone & Wolf-

erseberger (2012) suggested that although pricing is an important showrooming aspect, store associates and in-store assisting technologies play important role as well. Correspondingly, Rapp et al. (2015) elaborated on the relationship between showrooming and the salesperson and found out that retailers should invest in salesperson-consumer interaction through specific strategies and behaviours. In our case, we selected four criteria for consumers in order to purchase from a physical store, based on the previous showrooming-related factors: conventional store atmosphere, online store atmosphere, service support by salespeople utilizing sales supporting electronic technologies and a store's multichannel integration in order to create a seamless shopping experience. Store atmosphere in both offline & online variants was included, since in our case we investigate the omniretailing environment. The effect of retail salesperson was empowered with electronic technologies, in order to test omniretailing effects to him, too. It should be noted that although salespeople could be regarded as part of the conventional store atmosphere (human factor), in our case we examine them separately. The reason is twofold: to test human (e.g. personal selling techniques) vs environmental atmospheric effects and to separately examine the combination of human-technology effects on consumers' preferences. Omnichannel effects to showrooming were also incorporated as a criterion, based on its definition: a multichannel integration in order to create a seamless shopping experience (Levy, et al., 2013, p.67). As a result, our hypothesis is formulated as follows:

H4: *There are statistically significant differences between in-store internet users with different levels of showrooming intention, in terms of the importance they attach to conventional (H4.1) & online store atmosphere (H4.2), salespeople (H4.3) & omnichannel integration (H4.4) criteria in order to purchase from the physical store.*

Finally, it would be intriguing to discover if there are any differences regarding retailing mobile app and non app shoppers in relation to the previous offline purchase intention criteria. In other words, if retail mobile app assisted in-store internet shoppers attach more importance to each of these criteria in order to purchase from the physical store, in relation to non mobile app assisted in-store internet shoppers. That could be attributed to the enhanced mobile app UI and features, which could make these users to differ in terms of the previous criteria in relation to the others. Therefore, the following hypothesis could be originated:

H5: *There are statistically significant differences between mobile app and non mobile app assisted in-store internet shoppers, in terms of the importance they attach to conventional (H5.1) & online store atmosphere (H5.2), salespeople (H5.3) & omnichannel integration (H5.4) criteria in order to purchase from a physical store.*

For testing the research hypotheses, the study employs an exploratory quantitative empirical research design that took place in Greece in November 2014, in the context of an annual ELTRUN - The E-Business Research Center eCommerce survey. The data collection instrument of the national survey was an online questionnaire which received 815 valid answers from Internet users. The questionnaire was created in the Google forms platform and internet users were invited to participate via e-mail campaigns, display banners on popular Greek news sites & e-shops and social media. Questions included frequency of internet use at various channels, retailing-assisting mobile app utilization, as well as questions regarding 18 omniretailing practices & technologies within the physical store. These omniretailing practices & technologies were sorted according to the shopping process encounter, i.e. from the store entrance to the store checkout. Finally, they were asked about the importance they attach to the four aforementioned criteria, in order to purchase from a physical store. Statistical analysis was performed using SPSS version 20, and its outputs are presented and discussed at the following sections.

3 Findings And Discussion

Descriptive statistics confirm the forecast that was made back in 2011 that mobile internet will surpass desktop internet usage by 2014 (Wisniewski, 2011): 86% choose mobile phones for internet utiliza-

tion, whereas 78% use desktops. It was also found that 80% of retailing mobile app users use internet in-store and 46% of them do it often. What’s more, 70% of in-store internet users use retailing mobile apps in order to assist their purchases & 56% of them attach great importance to in-store retail-assisting mobile app or mobile sites. Besides, 39% use them often, whereas 31% rarely. It seems obvious that they use them along with internet inside physical stores, in order to facilitate shopping, and 60% of them respond that they attach great importance to them. Our initial research question is partially answered by the previous descriptive statistics. Retailing mobile app consumers definitely utilize them in-store. However, in order to validate hypothesis H1 additional statistical tests should be applied. Specifically, ANOVA was performed in order to test whether there are significant statistical differences between shoppers with different frequency of retailing mobile apps use, in terms of their in-store internet use. Shoppers were separated into three groups in order to perform the test: Group 1: Non-app users, Group 2: Rare mobile retailing app users, Group 3: Frequent mobile retailing app users. Frequency of in-store internet use was measured on a 5-point likert scale. Descriptive statistics of these groups are shown in Table 1.

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|----------|-----|-------|----------------|------------|----------------------------------|-------------|---------|---------|
| | | | | | Lower Bound | Upper Bound | | |
| Non | 264 | 2,080 | ,9419 | ,0580 | 1,965 | 2,194 | 1,0 | 4,0 |
| Rare | 285 | 2,102 | ,9606 | ,0569 | 1,990 | 2,214 | 1,0 | 4,0 |
| Frequent | 266 | 2,733 | ,8469 | ,0519 | 2,631 | 2,835 | 1,0 | 4,0 |
| Total | 815 | 2,301 | ,9658 | ,0338 | 2,234 | 2,367 | 1,0 | 4,0 |

Table 1: Descriptives of the Frequency of Internet use inside Physical Stores

ANOVA results (Table 2) reveal that the null hypothesis is rejected (p value < .05), and that significant statistical differences exist only between frequent user group and all the others (Table 3).

| | Sum of Squares | df | Mean Square | F | Sig. |
|----------------|----------------|-----|-------------|--------|------|
| Between Groups | 73,922 | 2 | 36,961 | 43,786 | ,000 |
| Within Groups | 685,428 | 812 | ,844 | | |
| Total | 759,350 | 814 | | | |

Table 2: ANOVA for Groups of Different Frequency of Internet use inside Physical Stores

| Dependent Variable: Frequency of Internet use inside Physical Stores - Tukey HSD | | | | | | |
|--|--|-----------------------|------------|------|-------------------------|-------------|
| (I) Frequency of Mobile Retailing Apps use | (J) Frequency of Mobile Retailing Apps use | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
| | | | | | Lower Bound | Upper Bound |
| Non | Rare | -,0222 | ,0785 | ,957 | -,206 | ,162 |
| | Frequent | -,6535 | ,0798 | ,000 | -,841 | -,466 |
| Rare | Non | ,0222 | ,0785 | ,957 | -,162 | ,206 |
| | Frequent | -,6313 | ,0783 | ,000 | -,815 | -,447 |
| Frequent | Non | ,6535 | ,0798 | ,000 | ,466 | ,841 |
| | Rare | ,6313 | ,0783 | ,000 | ,447 | ,815 |

Table 3: ANOVA Post Hoc Multiple Comparisons Test

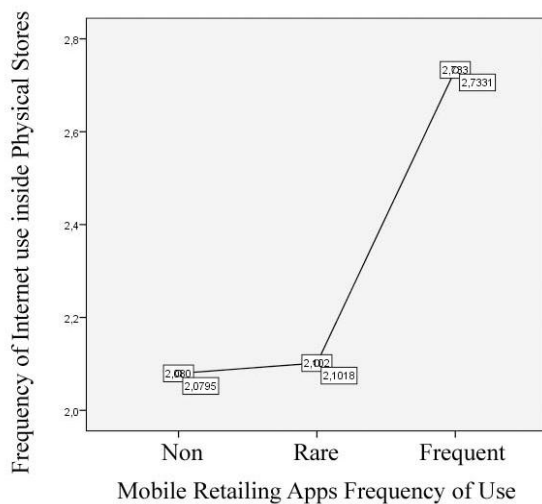


Figure 1: Frequency of Mobile Retailing Apps - In-Store Internet Use Plot

Drawing from the ANOVA results and also Figure 1 (Frequency of Mobile Retailing Apps - In-Store Internet Use Plot) we predict a positive correlation between mobile apps use and in-store internet use, which is statistically significant. Therefore, we validate our hypothesis by performing a correlation test. A Pearson product-moment correlation was run (Table 4) to determine the relationship between shoppers’ mobile apps use and their in-store internet use, which was found to be significant ($r = .273$, $n = 815$, $p < .05$). Consequently, hypothesis H1 is accepted.

| | | Frequency of Internet use inside Physical Stores | Frequency of Mobile Retailing Apps use |
|--|---------------------|--|--|
| Frequency of Internet use inside Physical Stores | Pearson Correlation | 1 | ,273 |
| | Sig. (2-tailed) | | ,000 |
| | N | 835 | 815 |
| Frequency of Mobile Retailing Apps use | Pearson Correlation | ,273 | 1 |
| | Sig. (2-tailed) | ,000 | |
| | N | 815 | 965 |

Table 4: Pearson Correlation for Frequency of Mobile Retailing Apps - In-Store Internet Use

In order to test hypothesis H2, we compared the 18 online practices & technologies inside the physical stores separating them at 18 groups, while performing ANOVA between them, in terms of the importance that in-store Internet retail mobile app assisted users attach to them. In this way not only we will grade them in terms of mean scores, but also we can identify significant statistical differences between them.

| | Online practices & technologies applied inside physical stores | Mobile Retailing App Consumers Means | Non App Consumers Means |
|----------------------------------|--|--------------------------------------|-------------------------|
| Importance attached by consumers | 16. Fast electronic checkouts without queues | 4,27 | 3,96 |
| | 3. Ability to buy in-store with internet prices, as a result of an electronic check-in in the physical store | 4,20 | 3,99 |
| | 6. In-store price comparison, which could lead to showrooming | 4,09 | 4,11 |
| | 1. Free in-store wifi | 3,91 | 3,64 |
| | 11. Product stock electronic availability | 3,89 | 3,83 |

| | | |
|--|------|------|
| 15. Product electronic search & map navigation to them | 3,82 | 3,75 |
| 4. Special prices, coupons, offer alerts at the store's entrance | 3,80 | 3,93 |
| 12. Loyalty points electronic access | 3,67 | 4,01 |
| 14. In-store location-based offers | 3,59 | 3,76 |
| 8. Access to user opinions, product presentations & reviews | 3,57 | 3,28 |
| 18. In-store retail-assisting mobile site (accessible via wifi) or mobile app | 3,55 | 3,53 |
| 17. Mobile payments | 3,52 | 3,11 |
| 5. Access to electronic profile & purchase history | 3,46 | 2,95 |
| 13. Self-service assisting technologies | 3,43 | 3,17 |
| 7. Electronic recommender systems | 3,28 | 2,51 |
| 9. Product/service posts and comments on social networks | 3,13 | 2,65 |
| 10. Email send & receive | 3,06 | 2,51 |
| 2. Electronic check-in in the physical store (e.g. via wifi, foursquare, swarm, facebook, etc) | 2,83 | 2,51 |

Table 5: Mean Ranking of Importance attached to Online Practices & Technologies Applied In-Store by Mobile Retailing App Consumers Vs Non App Consumers

The ranking of these practices and technologies according to their mean scores is presented in Table 5. The three most preferred ones are “Fast electronic checkouts without queues”, the “Ability to buy in-store with internet prices, as a result of an electronic check-in in the physical store” and “In-store price comparison, which could lead to showrooming”. It should be noted that they are the only ones with mean scores above 4 in the 5-point likert scale and that ANOVA post-hoc comparison showed that there are no significant statistical differences between them. In contrast, significant statistical differences do exist between these three practices and all the rest. The “Ability to buy in-store with internet prices, as a result of an electronic check-in in the physical store” also depicts consumers’ price sensitivity, which strikingly elevates the “Electronic check-in in the physical store” feature from the last place (score: 2,83) to the second one (score: 4,20). In addition to this, we calculated the percentage of consumers that attach great importance (over 4 points at the 5-point likert scale) to the previous practices. It appears that “In-store price comparison, which could lead to showrooming” now comes first, surpassing the other two (Table 6). All these results advocate that hypothesis H2 is accepted.

Next, we aim at exploring hypothesis H3, that is whether the previous results differ in terms of whether the users utilize mobile apps for retailing or not. Our test sample consisted only in-store internet users, therefore it would be interesting to explore if web-only users have the same technology preferences with mobile app users. Table 5 depicts their preferences (mean scores) by comparison. We performed an independent samples t-test which showed that null hypothesis is rejected for 12 sub-hypotheses. As a result there are significant statistical differences between mobile app and non mobile app assisted in-store internet shoppers, in terms of the importance they attach to these 12 online practices & technologies applied inside physical stores. “In-store price comparison, which could lead to showrooming” practice ranks first among non retailing mobile app consumers and wasn’t among the 12 ones supported by our hypotheses. Overall, the practices that didn’t show significant statistical differences, and thus the related sub-hypotheses were rejected, were the following:

1. In-store price comparison, which could lead to showrooming
2. Product stock electronic availability
3. Product electronic search & map navigation to them
4. Special prices, coupons, offer alerts at the store's entrance
5. In-store location-based offers
6. In-store retail-assisting mobile site (accessible via wifi) or mobile app

Therefore, these practices are considered by both retail mobile app and non retail app customers to be of equal high importance. Furthermore, half of these technologies feature location-based services (#3, #4, #5). It seems that the attached importance to these services is equal to both user groups, although only mobile app consumers have full access to them (e.g. iBeacon, gps). This finding shows that these technologies should be implemented into both technology approaches (#6 verifies that, too). On the other hand, there is low availability of these types of apps in the application stores. Therefore, non app consumers may utilize apps if retailing apps that lie in these categories are available, since they share the same preferences for them as the other mobile app consumer group.

| | Mobile Retailing App Consumers Percentage | Non App Consumers Percentage | Percentage Differences |
|--|---|------------------------------|------------------------|
| 6. In-store price comparison, which could lead to showrooming | 83% | 89% | -6% |
| 16. Fast electronic checkouts without queues | 81% | 67% | 14% |
| 3. Ability to buy in-store with internet prices, as a result of an electronic check-in in the physical store | 76% | 78% | -1% |
| 11. Product stock electronic availability | 71% | 71% | 0% |
| 4. Special prices, coupons, offer alerts at the store's entrance | 66% | 76% | -10% |
| 8. Access to user opinions, product presentations & reviews | 63% | 44% | 18% |
| 1. Free in-store wifi | 62% | 53% | 9% |
| 12. Loyalty points electronic access | 60% | 67% | -6% |
| 15. Product electronic search & map navigation to them | 59% | 63% | -4% |
| 18. In-store retail-assisting mobile site (accessible via wifi) or mobile app | 56% | 63% | -7% |
| 14. In-store location-based offers | 54% | 58% | -5% |
| 17. Mobile payments | 51% | 39% | 12% |
| 5. Access to electronic profile & purchase history | 49% | 39% | 10% |
| 13. Self-service assisting technologies | 49% | 34% | 15% |
| 7. Electronic recommender systems | 45% | 22% | 22% |
| 10. Email send & receive | 38% | 6% | 32% |
| 9. Product/service posts and comments on social networks | 35% | 23% | 12% |
| 2. Electronic check-in in the physical store (e.g. via wifi, foursquare, swarm, facebook, etc) | 33% | 22% | 11% |

Table 6: Percentage of Consumers that attach great importance to the Online Practices & Technologies Applied In-Store

Finally, we calculated the percentage of non retail app users that regarded each technology of utmost importance (4 & 5 in the 5-point likert scale of preference). At Table 6 we rank these preferences, in comparison with mobile app users. In-store price comparison, which could lead to showrooming gathers the highest percentage of the sample that consider it to be of utmost importance, highest than retailer mobile app users (89% vs 83%). In addition to this, non app consumers score higher than mobile app ones regarding "In-store price comparison, which could lead to showrooming" in the mean scores (Table 5), which is striking since mobile apps feature easier price comparison techniques, e.g. through camera barcode recognition. That could be attributed to either low performance of mobile apps in this

category (e.g. troublesome barcode recognition) or higher desire for appropriate apps by non app consumers because they do not have them available.

Finally, we observe that “Email send & receive” & “Electronic recommender systems” are the two technologies with the highest differences in percentages of users that regarded each technology of utmost importance (32% & 22% respectively). The differences are in favour of retailing mobile app consumers, which, in the case of email, indicates that these users attach more importance to checking emails via apps in-store than the others, since apps provide push mechanism through the OS, which is more efficient. However, email activities rank 17th at our standings. The higher percentage of user preference to “Electronic recommender systems” probably indicates that it is a feature best implemented through apps, since it involves more complicated functionalities and UI.

Next, drawing from our results regarding in-store price comparison, which could lead to showrooming, we aim at testing hypothesis H4 regarding showrooming intention. For this purpose we performed an independent samples t-test between the respondents group that attach high importance to in-store price comparison, which could lead to showrooming and those that don't. Results show that in-store internet users that attach great significance to in-store price comparison, which could lead them to showrooming, consider service support by salespeople utilizing sales supporting electronic technologies (H4.3) and a store's multichannel integration in order to create a seamless shopping experience (H4.4) more important than those that don't attach great significance to it (Table 8). Salespeople utilizing sales supporting electronic technologies is considered to be the most important (Table 7). In contrast, there are no statistically significant differences between these consumer groups in terms of the importance they attach to online & offline store atmosphere in order to purchase from a physical store (Table 8). Therefore, sub-hypotheses H4.1 & H4.2 are rejected.

| | In-store price comparison, which could lead to showrooming | N | Mean | Std. Deviation | Std. Error Mean |
|---|--|-----|-------|----------------|-----------------|
| Store's conventional atmosphere | >= 4,0 | 523 | 3,780 | 1,1225 | ,0491 |
| | < 4,0 | 78 | 3,885 | ,6026 | ,0682 |
| Service support by salespeople utilizing sales supporting electronic technologies | >= 4,0 | 523 | 4,340 | ,9091 | ,0398 |
| | < 4,0 | 88 | 3,886 | ,8767 | ,0935 |
| Online store's atmosphere | >= 4,0 | 523 | 3,975 | 1,1130 | ,0487 |
| | < 4,0 | 88 | 4,000 | ,6781 | ,0723 |
| Multichannel integration in order to create a seamless shopping experience | >= 4,0 | 523 | 4,036 | ,9926 | ,0434 |
| | < 4,0 | 78 | 3,756 | ,9828 | ,1113 |

Table 7: Descriptives of offline purchase intention criteria between in-store internet users with different levels of showrooming intention

| | | t-test for Equality of Means | | | | | | |
|---------------------------------|----------------------|------------------------------|---------|-----------------|-----------------|-----------------------|---|-------|
| | | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | Lower | Upper |
| Store's conventional atmosphere | Eq. var. assumed | -,805 | 599 | ,421 | -,1045 | ,1299 | -,3595 | ,1505 |
| | Eq. var. not assumed | -1,243 | 170,572 | ,215 | -,1045 | ,0841 | -,2704 | ,0614 |
| Service support | Eq. var. assumed | 4,356 | 609 | ,000 | ,4540 | ,1042 | ,2493 | ,6587 |

| | | | | | | | | |
|--|----------------------|-------|---------|------|--------|-------|--------|-------|
| by salespeople utilizing sales supporting electronic technologies | Eq. var. not assumed | 4,470 | 120,674 | ,000 | ,4540 | ,1016 | ,2529 | ,6550 |
| Online store's atmosphere | Eq. var. assumed | -,203 | 609 | ,839 | -,0249 | ,1223 | -,2651 | ,2154 |
| | Eq. var. not assumed | -,285 | 177,683 | ,776 | -,0249 | ,0871 | -,1968 | ,1471 |
| Multichannel integration in order to create a seamless shopping experience | Eq. var. assumed | 2,326 | 599 | ,020 | ,2799 | ,1203 | ,0436 | ,5162 |
| | Eq. var. not assumed | 2,344 | 101,863 | ,021 | ,2799 | ,1194 | ,0430 | ,5168 |

Table 8: T-Test between in-store internet users with different levels of showrooming intention regarding offline purchase intention criteria

We also performed Pearson's correlation for each of the aforementioned criteria, regarding showrooming intention. Results reveal that there is a positive correlation between showrooming intention and importance attached to service support by salespeople utilizing sales supporting electronic technologies in order to purchase from a physical store, which is statistically significant ($r = .244$, $n = 611$, $p < .05$). In addition, there is a positive correlation between showrooming intention and importance attached to a store's multichannel integration in order to create a seamless shopping experience in order to purchase from a physical store ($r = .187$, $n = 611$, $p < .05$). On the contrary, there is no positive correlation between showrooming intention and importance attached to a store's online ($r = .033$, $n = 611$, $p=0.410$) & offline ($r = .043$, $n = 611$, $p=0.292$) store atmosphere in order to purchase from a physical store.

Last but not least, regarding hypothesis H5, we perform an independent samples t-test between mobile retailing app users and non app users regarding the same criteria of hypothesis H4. It turns out that there are statistically significant differences between mobile app and non mobile app assisted in-store internet shoppers, only in terms of the importance they attach to online store atmosphere (H5.2) in order to purchase offline (Table 9). This result probably indicates that due to the additional and superior UI that mobile app assisted in-store internet shoppers interact, they respond more to online atmospherics, in order to purchase from the physical store. Thus, only sub-hypothesis H5.2 is accepted.

| | | t-test for Equality of Means | | | | | | |
|---|----------------------|------------------------------|---------|-----------------|-----------------|-----------------------|---|-------|
| | | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | Lower | Upper |
| Store's conventional atmosphere | Eq. var. assumed | ,302 | 608 | ,763 | ,0289 | ,0957 | -,1590 | ,2168 |
| | Eq. var. not assumed | ,279 | 279,119 | ,780 | ,0289 | ,1036 | -,1750 | ,2328 |
| Service support by salespeople utilizing sales supporting electronic technologies | Eq. var. assumed | ,952 | 618 | ,341 | ,0774 | ,0813 | -,0822 | ,2369 |
| | Eq. var. not assumed | ,921 | 305,154 | ,358 | ,0774 | ,0840 | -,0880 | ,2427 |
| Online store's atmosphere | Eq. var. assumed | 3,923 | 609 | ,000 | ,3663 | ,0934 | ,1830 | ,5497 |
| | Eq. var. not assumed | 3,355 | 246,478 | ,001 | ,3663 | ,1092 | ,1513 | ,5814 |

| | | | | | | | | |
|--|----------------------|------|---------|------|-------|-------|--------|-------|
| Multichannel integration in order to create a seamless shopping experience | Eq. var. assumed | ,135 | 608 | ,893 | ,0121 | ,0903 | -,1651 | ,1894 |
| | Eq. var. not assumed | ,136 | 309,261 | ,892 | ,0121 | ,0891 | -,1632 | ,1875 |

Table 9: T-Test between Mobile Retailing App Consumers and Non App Consumers regarding offline purchase intention criteria

4 Implications for Retail Entrepreneurs

The present study provides several managerial implications for mobile app entrepreneurs, as well as retail managers alike, towards developing the future retail store. Retailing mobile apps that provide consumer assisting features are still at their infancy. Mobile app developers should focus not only providing a supplemental shopping assistant interface, but also integrate this interface with hardware features that blend the physical and the online world seamlessly. In that way, they could provide the best of the two worlds, creating a superior shopping environment that could deter showrooming and provide added-value services. It should be noted that some retailers have already leveraged mobile apps as omniretailing assisting technologies either by providing in-store location based services & promotions (e.g. Apple stores app, Macy's Shopbeacon, Carrefour China app), loyalty points & social media integration (Guess Mobile app) or augmented reality support (American Apparel).

Observing Tables 5&6 we can extract several guidelines regarding the features that omniretailing mobile apps should offer. More specifically, they should offer deep integration with backend IS and POS systems in order to facilitate fast checkouts, unified pricing and realtime stock availability. Additional technologies such as in-store location based services are second runners, but shouldn't be neglected, either.

In addition to this, it seems that mobile apps could prove to be even more beneficial for store associates. Consumers value the salespeople-technology combination the most, therefore mobile apps could empower employees in a more powerful fashion. In that case, apps could be more effective by utilizing them on tablets, in order to provide a more spacious UI. Regarding, showrooming avoidance, literature also shows that specific personal selling techniques & strategies should also be adopted by salespeople, since technology on its own is not enough (Rapp et al., 2015). Therefore, retailers should invest on their human capital, while transforming into omniretailers, embracing omnichannel retailing principles and guidelines. Taking into consideration Table 6, we should advice mobile app entrepreneurs to offer anti-showrooming services for salespeople's mobile devices. Indicatively, these could include price comparison and price matching functionalities. Towards these directions, new apps, mainly for salespeople tablets, start to emerge (e.g. Shopkeep, Entersoft Mobile Retail Sales Assistant).

In sum, the future retail store should offer deep omnichannel integration, providing a 360 degree view of the customers (e.g. incorporating universal analytics), unifying the offline and the online shopping experience. To that end, new omniretailing software platforms were recently introduced aiming at merging online & offline operations providing universal analytics (e.g. Euclid Analytics, Index, RetailNext, Prism). This integration could additionally be assisted with the use of apps, but stores should be also enhanced with supplemental technologies that offer location-based services (e.g. iBeacon), efficient & beneficial electronic check-in for consumers, as well as fast electronic checkouts without queues. Towards the last direction, the store could support mobile payments, or even eliminate checkouts completely. In vision of that, a recent Amazon patent (Amazon, 2015), employing RFID technology and ubiquitous video cameras, shows that the online retailer may attempt to disrupt the physical

retail domain towards that direction by opening bricks ‘n’ mortar stores³ that offer automated check-outs⁴.

5 Conclusions, Limitations and Future Research

As E-Commerce practices are desired by consumers in the physical stores, retailing mobile apps seem to play an important role in this behaviour, integrating retail channels. Increased use of these apps has been found to take place in-stores, accompanied with increased use of mobile internet. Price-centric apps dominate users’ preferences with “In-store price comparison, which could lead to showrooming” to gather the largest percentage of them that regard it of utmost importance. Hence, showrooming intention is high among in-store internet users, both retailing mobile app and non app ones. Interestingly, for that target consumer group (high showrooming intent), service support by salespeople utilizing sales supporting electronic technologies and omnichannel integration were found to be regarded as more important than the group that didn’t care about showrooming. This finding leads us to believe that apart from price-matching strategies (since these consumers seem to be price-centric), increased importance should be placed at the role of salespeople in the physical store, as well as at omnichannel integration strategies.

Regarding the role of salespeople, the results are consistent with related studies (Zhang & Oh, 2013; Monteleone & Wolferseberger, 2012; Rapp et al., 2015; Pantano & Viassone, 2015) which emphasized on the dominant role that store associates play, coping with that emerging consumer behaviour. In particular, Zhang & Oh (2013) stressed on the role of service support, Monteleone & Wolferseberger (2012) on salespeople assisting technologies, Rapp et al. (2015) on salesperson-consumer interaction and Pantano & Viassone (2015) on service quality perception as an outcome of technology and/or salesperson interaction. Therefore, our criterion of “Service support by salespeople utilizing sales supporting electronic technologies” is validated as a means of battling showrooming, since users that tend to engage in such behaviour attach significantly more importance to it (more than any other criteria) in order to purchase from the physical store that they have visited.

As far as the omnichannel integration criterion in concerned, that is, a store’s multichannel integration in order to provide a seamless shopping experience, Pantano & Viassone (2015) provided empirical evidence that it can prevent showrooming and suggested the use of channel integrating technologies to accomplish it (iBeacon, mobile apps and smartphones). This finding is also consistent with our statistical findings regarding these technologies that gather increased attention by consumers. In addition to this, Zhang & Oh (2013) also suggested that cross-channel services lead to customer retain. Nevertheless, Chiu et al. (2011) found that multichannel self-efficacy positively affects showrooming; therefore multichannel integration should be carried out cautiously. On the other hand, store atmosphere, both in conventional and online variants, though considered important, it doesn’t attract consumers with showrooming intention more than the others. However, these results are not consistent with Shukla & Babin (2013), Pantano & Viassone (2015) and Chiu et al. (2011) findings indicating that store atmosphere affects showrooming. On the contrary, they are in line with Heitz-Spahn (2013) claims that channel aesthetics as components of store atmosphere do not influence cross-channel free-riding behaviour. However, they propose that this behaviour could be fought with appropriate mobile applications.

3 <http://www.theguardian.com/technology/2015/feb/03/amazons-first-store-opens-indiana>

4 http://www.retailwire.com/discussion/18195/could-amazons-brick-and-mortar-invention-eliminate-checkout-lines?utm_campaign=RW%20Discussions&utm_content=13995290&utm_medium=social&utm_source=twitter

Captivatingly, retailing mobile apps in-store internet users regard a store's online store atmosphere of utmost importance in order to purchase from that store's physical counterpart. The importance mean score they attach is even higher than the store's conventional store atmosphere equivalent. What's more, among this consumer group and those that don't use retailing mobile apps, there were found to be statistical differences only regarding the online store atmosphere's importance in order to purchase offline. As a result, that could mean that proper mobile app atmospherics could also influence retailing mobile app consumers' showrooming intention in-store.

The study encloses several limitations that are mainly attributed to the research setting and method. First of all, our sample consisted of solely internet users and, therefore, our results cannot be generalised to the whole population, who may not be interested in online practices within physical stores. In addition to this, consumers were asked which practices & technologies they considered most important in-store and not which they actually employ. The reason for that was the availability of most practices & technologies, which were too advanced at that time for stores to support them, especially in the form of mobile apps. For that reason, consumers could respond differently if they had actual experience of them in the conventional shopping environment. Last but not least, to the best of our knowledge, the store atmosphere notion has not been transferred to the mobile apps domain. While web atmospherics components could be applied to mobile web one with little modifications (Manganari et al., 2007), mobile apps, as discussed, provide features unique to the online world that may influence our online atmosphere-related results.

In order to verify and expand our findings, researchers are encouraged to employ experimental design approaches in real physical stores, in order to test omniretailing effects in practice. Field experiments should definitely exploit the use of retailing mobile apps, since they are the most suitable choice towards blending physical with virtual experiences. Also, the interplay of multiple atmospheric cues, both online & offline, through omnichannel integration remain unexplored. Hence, it would be intriguing to explore consumer behaviour and the showrooming phenomenon specifically, inside the future retail store, where the Omnichannel Retailing Store Atmosphere is present.

References

- Adobe. (2013). *Mobile shopping: retail app usage on the rise*. Retrieved from <http://blogs.adobe.com/digitalpublishing/2013/01/mobile-shopping-retail-app-usage-on-the-rise.html>
- Amazon. (2015). *Transitioning items from a materials handling facility*. Retrieved from <https://www.google.com/patents/US20150012396>
- Agrebi, S., & Jallais, J. (2015). "Explain the intention to use smartphones for mobile shopping." *Journal of Retailing and Consumer Services*, 22, 16–23.
- Burke, R. R. (2002). "Technology and the Customer Interface: What Consumers Want in the Physical and Virtual Store." *Journal of the Academy of Marketing Science*, 30(4), 411–432.
- Charland, A., & Leroux, B. (2011). "Mobile application development." *Communications of the ACM*, 54(5), 49.
- Chiu, H.-C., Hsieh, Y.-C., Roan, J., Tseng, K.-J., & Hsieh, J.-K. (2011). "The challenge for multichannel services: Cross-channel free-riding behavior." *Electronic Commerce Research and Applications*, 10(2), 268–277.
- Dailey, L. (2004). "Navigational web atmospherics." *Journal of Business Research*, 57(7), 795–803.
- Eroglu, S. A., & Machleit, K. A. (1993). "Atmospheric Factors in the Retail Environment: Sights, Sounds and Smells." *Advances in Consumer Research*, 20(1), 34.
- Fairchild, A. M. (2014). "Extending the Network: Defining Product Delivery Partnering Preferences for Omni-channel Commerce." *Procedia Technology*, 16, 447–451.
- Gerpott, T. J., & Thomas, S. (2014). "Empirical research on mobile Internet usage: A meta-analysis of the literature." *Telecommunications Policy*, 38(3), 291–310.

- Heitz-Spahn, S. (2013). "Cross-channel free-riding consumer behavior in a multichannel environment: An investigation of shopping motives, sociodemographics and product categories." *Journal of Retailing and Consumer Services*, 20(6), 570–578.
- Huy, N. P., & vanThanh, D. (2012). "Evaluation of mobile app paradigms." In *Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia - MoMM '12* (p. 25). New York, New York, USA: ACM Press.
- Lazaris, C., Vrechopoulos, A., Doukidis, G., & Fraidaki, K. (2015). "The Interplay of Omniretailing & Store Atmosphere on Consumers' Purchase Intention towards the Physical Retail Store." In *12th European, Mediterranean & Middle Eastern Conference on Information Systems (EMCIS)*. Athens, Greece.
- Lazaris, C., Vrechopoulos, A., Fraidaki, K., & Doukidis, G. (2014). "Exploring the "Omnichannel" Shopper Behaviour." In *AMA SERVSIG, International Service Research Conference, 13-15 June*. Thessaloniki, Greece.
- Levy, M., Weitz, B., & Grewal, D. (2013). *Retailing Management*. McGraw-Hill/Irwin; 9th edition.
- Luo, Q., Oh, L.-B., Zhang, L., & Chen, J. (2014). "Examining the showrooming intention of mobile-assisted shoppers in a multichannel retailing environment." In *Pacific Asia Conference on Information Systems (PACIS)* (p. Paper 141).
- Magrath, V., & McCormick, H. (2013). "Marketing design elements of mobile fashion retail apps." *Journal of Fashion Marketing and Management: An International Journal*, 17(1), 115–134.
- Mahatanankoon, P., Wen, H. J., & Lim, B. (2005). "Consumer-based m-commerce: exploring consumer perception of mobile applications." *Computer Standards & Interfaces*, 27(4), 347–357.
- Manganari, E., Siomkos, G., & Vrechopoulos, A. (2007). "Atmospheric qualities in mobile commerce: an initial approach." In *Proceedings of the 36th European Marketing Academy Conference (EMAC)* (pp. 1–7). Reykjavik, Iceland: McGraw-Hill.
- Mikkonen, T., & Taivala, A. (2011). "Apps vs. open web: The battle of the decade." In *Proceedings of the 2nd Workshop on Software Engineering for Mobile Application Development* (pp. 22–26).
- Monteleone, P., & Wolferseberger, J. (2012). *Showrooming and the Price of Keeping Buyers In-Store*. *GroupM Next*. Retrieved from <https://www.scribd.com/doc/103349382/GroupM-Next-White-Paper-Showrooming-and-the-Price-of-Keeping-Buyers-In-Store#download>
- Ortis, I., & Casoli, A. (2009). Technology Selection: IDC Retail Insights Guide to Enabling Immersive Shopping Experiences. *IDC Retail Insights*. Retrieved from <http://idg.com/www/pr.nsf/ByID/IDGC-8MBQMJ>
- Pantano, E., & Viassone, M. (2015). "Engaging consumers on new integrated multichannel retail settings: Challenges for retailers." *Journal of Retailing and Consumer Services*, 25, 106–114.
- Parker, R., & Hand, L. (2009). *Satisfying the Omnichannel Consumers Whenever and Wherever They Shop*. *IDC Retail Insights*. IDC Retail Insights. Retrieved from <http://www.amazon.com/Satisfying-Omnichannel-Consumers-Whenever-Wherever/dp/B002PW0G46>
- Quint, M., Rogers, D., & Ferguson, R. (2013). *Showrooming and the rise of the mobile-assisted shopper*. *Columbia Business School - Aimia*. Retrieved from https://www4.gsb.columbia.edu/filemgr?file_id=7313935
- Rapp, A., Baker, T. L., Bachrach, D. G., Ogilvie, J., & Beitelspacher, L. S. (2015). "Perceived customer showrooming behavior and the effect on retail salesperson self-efficacy and performance." *Journal of Retailing*, 91(2), 358–369.
- Rigby, D. (2011). The Future of Shopping. *Harvard Business Review*, 89(12), 64–75.
- Shukla, P., & Babin, B. J. (2013). "Effects of consumer psychographics and store characteristics in influencing shopping value and store switching." *Journal of Consumer Behaviour*, 12(3), 194–203.
- Spence, C., & Gallace, A. (2011). "Multisensory design: Reaching out to touch the consumer." *Psychology and Marketing*, 28(3), 267–308.
- Verhoef, P. C., Kannan, P. K., & Inman, J. J. (2015). "From Multi-Channel Retailing to Omnichannel Retailing." *Journal of Retailing*, 91(2), 174–181.

- Willmott, B. (2014). "Retail showrooms, mobile sales." *Journal of Direct, Data and Digital Marketing Practice*, 15(3), 229–232.
- Wisniewski, J. (2011). "Mobile That Works for Your Library." *Online*, 35(1), 54–57.
- Wurmser, Y. (2014). *Retailers Playing Catch-Up with Consumers*. eMarketer. Retrieved from http://www.emarketer.com/public_media/docs/state_of_omnichannel_retail.pdf
- Xu, J., Forman, C., Kim, J. B., & Van Ittersum, K. (2014). "News media channels: Complements or substitutes? Evidence from mobile phone usage." *Journal of Marketing*, 78(4), 97–112.
- Zhang, L., & Oh, L.-B. (2013). "Determinants of Multichannel Consumer Switching Behavior: A Comparative Analysis of Search and Experience Products." In *WHICEB 2013 Proceedings* (pp. 205–212). Wuhan, China.
- Zhao, Z., & Balagué, C. (2015). "Designing branded mobile apps: Fundamentals and recommendations." *Business Horizons*, 58(3), 305–315.

CHALLENGES TO NEWSPAPERS IN DIGITAL PLATFORM ECOSYSTEMS

Åkesson, Maria, Halmstad University, Halmstad, Sweden, maria.akesson@hh.se

Thomsen, Michel, Halmstad University, Halmstad, Sweden, michel.thomsen@hh.se

Abstract

Incumbent content providers such as newspapers experience radical changes in their business environment. Digital platforms are becoming increasingly important to innovate digital media services and business. The generativity of digital platforms afforded by digital technology offers new opportunities, but also new challenges. Digital platforms exist in an ecosystem of complex networks of actors and resources. In this paper we shed light on the challenges for incumbent content providers to adapt to and realize business opportunities in digital platforms. We do this from the viewpoints of newspaper management, platform provider and analysts of media business.

Keywords: Incumbent content providers, Digital platforms, Ecosystems, Generativity

1 Introduction

Newspapers of today are actors in a rapidly changing ecosystem. From having had a dominant position on the news market for more than 150 years, newspapers are now challenged by the digital economy. New digital technology, new media consumption patterns as well as new advertising models are changes that have challenged newspapers to innovate in digital media (Conboy and Steel, 2008; Picard, 2006). Undeniably, it has been very challenging to innovate news services, business models and adapt to the ecosystems that enable profitable business in digital media. Significant uncertainty exists related to the ecosystems of newspapers in digital environments. This uncertainty has even started a debate about the survival of traditional newspapers. Even so, it can be noted that up until now no new media has replaced another in the newspaper industry. That is, digital innovations adopted by newspaper organizations have not been disruptive in the meaning that they have replaced existing technology, but rather disruptive to their ecosystems as acknowledged by Christensen and Davis (2006). New digital innovations have changed or created new ecosystem relationships and disrupted socio-technical frames for decisions and value creation.

For incumbent content providers, such as newspapers, it has been proven challenging to innovate in digital media (Åkesson, 2009). Still, there are not many studies that explicitly have studied generativity and control in digital platforms and the business implications for newspapers. To advance knowledge on how media content business can transform to realize digital platform opportunities, we need to inquire into the facets of digitalization. That is, into the consequences of existing business practices in digital media, into the characteristics of digital media platforms, and into the business environment in digital ecosystems.

The objective of this study is to explore and shed light on the challenges for incumbent firms to adapt to and utilize business opportunities in digital platforms. More specifically, we have studied how newspaper organizations relate to challenges in digital platforms, and how platform actors relate to business practices of newspaper organizations. Although there is research suggesting that leveraging innovation in digital platforms requires balancing generativity and control (Eaton et al., 2015; Förderer et al., 2014), there are still few studies that have focused on the relationship between incumbent content providers and other platform actors. Further, there is a lack of research into the strategizing underlying decisions taken by ecosystem actors (Ghazawneh and Henfridsson, 2015). It can be assumed that

the rationality behind a new digital platform provider and incumbent content providers are different. However, we know little about what characterizes the challenges associated with decisions, and how the rationale behind decisions differ between platform ecosystem actors (Ghazawneh and Henfridsson, 2015). We therefore address the following overall research question: *How are incumbent content providers challenged in digital platform ecosystems?*

In this paper, we report from an ongoing research project in collaboration with the newspaper industry. The project is aimed at developing new insights on how the digital media ecosystem is changing, and how the newspaper industry is challenged by this change. Theoretically, we draw on previous research on generativity and digital platforms (see e.g. Yoo, 2013; Tilson et al., 2010) to understand how newspapers are challenged by digitalization of business. Empirically, we illustrate how newspaper management reason about digital media platforms and how other ecosystem actors construe newspapers practice. Finally, we interpret and discuss the results in the light of digital platform generativity (see e.g. Eaton et al., 2011).

2 Related literature

2.1 Background to changes in newspaper industry

The newspaper industry has been very profitable over time compared to other industries (Picard, 2003). Since the Second World War the print newspaper market has been mature and apart from evening press very few new newspapers have started or shut down since then. In other words, newspaper industry has up until recently existed in a stable and undisturbed ecosystem. However, during the latest decades, newspapers all over the world are suffering from decreasing print circulation and declining advertising revenues as readers and advertisers are turning to digital media (Picard, 2006; Stone et al., 2012). Even though newspaper services have been present on the Internet since the mid 90's and in mobile phone platforms since the end of the 90's (Boczkowski, 2004), newspapers have experienced difficulties to build digital business. The traditional business models of newspapers have not worked very well in digital media and the competition has been difficult to meet (Ihlström Eriksson et al., 2008). This situation has provoked a great need for innovation in the newspaper industry (Küng, 2008; Åkesson, 2009).

Driven by the shifting customer and advertiser behaviors, the media landscape is changing at an unprecedented pace. Media consumption is now less restricted to particular time and places, it is differentiated between different brands and media (Stone et al., 2012), and increasingly social and mobile. Indeed, people's increased physical mobility and the on-going diffusion and adoption of new mobile technologies are gradually transforming the way media is consumed. These changes are for example mirrored in decreasing print newspaper circulation and the radical increase in mobile news consumption (Stone et al., 2012). Advertisers are increasingly turning to digital media and they have need for more targeted and accurate advertising (Åkesson and Ihlström Eriksson, 2009). Accordingly, many newspaper organizations are putting more effort into digital media to find new business opportunities.

The confluence of these shifts drives newspaper organizations to engage with diverse actors thus bringing newspapers into a new ecosystem with implications on strategies and business models. The news media ecosystem is evolving into a complex network of interacting actors providing a broad range of news services to end-consumers and advertisers. Previous research suggests that business models in complex ecosystems go beyond individual firms in complex value networks of actors. To develop successful business models in complex ecosystems, firms need to carefully orchestrate inter-firm relationships, maintain and develop core competencies, and take network position and value creation into account (Chesbrough, 2006; Åkesson, 2009). This includes assessing ecosystem risks in order to establish more realistic expectations and profitable business models (Adner, 2006). Newspapers are engaged in networks of relationships with, among others, newspapers, publication system providers, advertising agencies, advertisers, and consumers. The complexities of the ecosystem increase in multi-fold as new actors are emerging, new relations are formed, and the traditional distribution of power is

shifted. Digital innovation has transformed and widened the relationships of newspapers. Newspapers have not been engaged with for example telecom providers until the opportunity of offering mobile news services on mobile platforms emerged. New digital services such as mobile Internet and social media are changing the relation between newspaper and consumer (Stone et al., 2012).

This has led to incumbents' inertia for the exploration of digital business. Some scholars argue that the newspaper industry has been infatuated by historical success (see e.g. Boczkowski, 2005; Picard, 2006), they have not been managed to explore digital opportunities in the pace required to transform from print to digital business. O'Reilly and Tushman (2004) highlighted in a study of the newspaper USA Today, that newspaper organization's capability to exploit the opportunities afforded by digital publishing depends on its ability to reinvent itself in parallel with exploiting the incumbent print business. Christensen et al. (2012) provide similar conclusions based an analysis of the newspaper industry suggesting that the industry requires the re-inventing of organizations to survive. This digital force continuously disrupts the fundamentals underlying newspaper business models, and challenges the culture and core identity of newspaper organizations (Åkesson and Ihlström Eriksson, 2009). The current pace of digitalization is faster than anyone could anticipate just a few years back (Christensen et al., 2012).

2.2 Ecosystems of digital platforms

Digital platforms have emerged as a dominant model of digital service innovation. A digital platform is a system of digital resources that can be used, re-combined and extended by a network of partner firms (Tiawana et al., 2010). A platform provider or owner can be a firm or a network of firms that provides the platform as a marketplace for content such as applications and media content (Ghazawneh and Henfridsson, 2015).

Digital platforms provide an open, collaborative, networked, and complex business ecosystem. A digital platform ecosystem is a collective of organizations having a common interest in leveraging digital service innovation (Ghazawneh and Henfridsson, 2015). Digital platform ecosystems are gaining increasing importance for content providers as digital innovation environments in which they compete and collaborate at the same time. In a platform ecosystem, firms can create value that no single firm could leverage alone (Adner, 2006).

Digital platform ecosystems are open and generative. A platform owner depends on other actors in the ecosystem to leverage value and innovation (Tilson et al., 2010). Generativity is at the hart of the innovative opportunities that follow digital developments (Yoo et al., 2012; Tilson et al., 2010). The term refers to the socio-technical system that defines how organizational resources (Avital and Teeni, 2009), and technological resources (Zittrain, 2006) can lead to the emergence of new digital innovations (Yoo, 2013).

The foundational premise for the generativity of digital platforms is the unique characteristics of digital technology (Yoo, 2013). As outlined by Yoo et al., (2010), first, digital technology is characterized by homogenization by the use of binary digits for all types of data. Second, digital technology is re-programmable, and third, self-referential. These three characteristics enable a layered modular architecture separating hardware (devices), network technology, services and content enables relatively independent new re-combinations and innovation within and between layers. This architecture makes content independent of physical limitations such as devices and geographical place. Taking the example of Apple, new apps and functions can be added to an iPhone or an iPad and thus re-configure the services and content without any changes to the device as such.

In this paper, we refer to generativity as an attribute of digital platforms enabling the capacity to provide new and varied re-combinations and a key source of value (Tilson et al., 2010; Yoo, 2013). Leveraging generativity involves innovating business models and creating new markets. However, the generativity of digital platforms also provides low entry barriers for new innovators compared to analogue media. Prices and performance of digital technology are relatively low compared to the investments needed for start-ups in analogue media, such as for example printing and distribution of news-

papers. In digital ecosystems, incumbent media content providers compete with new digital start-ups engaging in digital media platforms. The low barrier to entry makes it possible for firms to engage with heterogeneous actors in the eco-system to leverage service innovations in digital media platforms (Eaton et al., 2011). To appropriate generativity, the platform owner enacts control at the same time as too much control can restrain generativity. This balance is key to leverage value in digital platforms (Ghazawneh and Henfridsson, 2013; Förderer et al., 2014). Controlling generativity is therefore a strategy to create competitive barriers for rival platforms. The paradoxical relationship of control and generativity is that control that reduces generativity seeking to create barriers for emerging innovations, can also stimulate generativity. Similarly, the stimulation of generativity that can result in emerging and unexpected innovations can also create the need for increased control (Eaton et al., 2011).

3 Research approach

The approach to this study was qualitative (Orlikowski and Baroudi, 1991; Walsham, 1995). It was oriented towards content provider assumptions, knowledge, and experiences of digital innovation efforts. The study was done in the context of an on-going research project named New(s) Media Ecosystem. The project is a collaborative initiative with Swedish newspaper organizations with the ambition to cope with the transition from print to digital publishing. The overall goal of the project is to explore how newspapers can innovate in digital platforms. The project started in 2013 and will be finalized in April 2016. The initial participating newspaper organizations are Aftonbladet, Expressen, Göteborgs-Posten, Helsingborgs Dagblad, Mediabolaget Västskusten AB, MittMedia and Västerbottens-Kuriren. Some of these organizations have merged during the project due to the turbulent market conditions in the newspaper industry. The research team in the project consists of three senior researchers and two PhD candidates.

Since year 2000, Swedish newspaper circulation has declined by 40%. Advertising revenue has declined by 10-15% per year. Digital advertising revenue has increased but still, with few exceptions, is only 10-15% of the total advertising revenue. Digital readership has increased but revenues from digital consumers is not by far compensating for the loss of print circulation (Nordicom statistics).

During the project we have done interviews, group interviews and workshops with newspaper partners. We have also interviewed other actors in the ecosystem such as advertising agencies, analytics, platform providers, and newspaper readers. The main informants for this study are the project leaders appointed in each newspaper organization, analysts of media business, and the market director of a platform provider. The informants are presented in Table 1.

| Informant id | Role | Organization | Data collection activity |
|--------------|------------------------------------|-----------------------|--------------------------------|
| NP1 | Chief business developer (m) | Newspaper 1 | 3 group discussions (3 h each) |
| NP2 | Private market director (f) | Newspaper 2 | 3 group discussions (3 h each) |
| NP3 | Head of digital development (m) | Newspaper 3 | 3 group discussions (3 h each) |
| NP4 | Head of development (m) | Newspaper 4 | 3 group discussions (3 h each) |
| NP5 | Digital director (f) | Newspaper 5 | 3 group discussions (3 h each) |
| NA | Chief of analytics (m) | Newspaper association | Interview (2 h) |
| A2 | Editor in chief trade magazine (m) | Analysts 1 | Interview (1 h) |
| A3 | Chief Analysts, CEO (f) | Analysts 2 | Interview (1.5 h) |
| PP | Market director (m) | Platform provider | Interview (1 h) |

Table 1. Overview of informants and data collection activities

The two authors collected the data reported in this paper. The group discussions with the project leaders from newspaper organizations were part of three different two-day meetings in the project. In the first group discussion we discussed future scenarios for newspaper content and business. In the discussions, the platform provider role came clear as having a key position in the ecosystems. Therefore, we conducted an interview with a platform provider. In the second group discussion with the project lead-

ers from newspaper organizations we presented and discussed results from the interview with the platform provider. The responses to the results guided the interviews with analysts. These interviews were semi structured and covered topics relating to incumbent content providers in the era of digital platform ecosystems. The analysts all have many years of experienced in the newspaper industry and have all been professionally employed by newspaper organizations prior to becoming analysts. Informant A1 is specialized on analytics of advertising in digital platforms. Informant A2 does research on the digital transformation of newspaper on a global scale. Finally, informant A3 is editor in chief for a trade magazine covering media business. In addition to these data collection activities we participated in two media conferences on the topic of digital business. At these conferences we interacted with different actors in digital media ecosystems.

All group discussions with the newspaper project leaders were recorded and partially transcribed. Small talk and topics not relevant for the inquiry were left out in the transcription. The interviews were all transcribed in full. The analysis of the empirical material for the purpose of this paper follows a descriptive and interpretive approach (Walsham, 2006). In the analysis, we treated the newspaper informants as one unit of data, the platform provider as one, and the analysts as a third. The overall directing question was how newspapers as incumbent content providers are challenged in digital media platform ecosystems. We first reviewed transcripts and notes of the collected data to identify statements by the three different types of informants relating to the opportunities and challenges for newspapers in digital platforms. From these statements we identified two themes of challenges for focused coding; identity crisis and balancing openness and control in digital platform ecosystems. Thereafter we compared and reflected on the different informant's accounts. Both researchers did the interpretations of the data material. The results are organized in three sections representing the three categories of informants. The interpretation and discussion is guided by the literature on digital platforms and generativity.

4 Results

4.1 The platform provider

The platform provider offers content providers of print media such as magazines, newspapers, trade magazines and comic books, to sell their content on the platform. The service is mainly offered to tablets and mobiles. The platform has customers in 100 countries. The revenue share is 30 % to the content provider and 70 % to the platform owner. The revenue model to readers is a fixed fee of 10 euros per month.

The platform provider predicts that media ecosystem will be under change for a long time ahead. One of the driving trends is platforms that package and bundle media content to the consumer convenience. There are according to the informant many examples such as Netflix, Spotify and Amazon. To leverage value, these platforms need to offer a variety of titles. The challenge is to build up enough content to be interesting to the media consumer. All these platforms have managed to pass that breaking point, and the informant regards their platform to have managed that with magazines. They have not managed to attract newspaper content providers to the platform:

We have a few newspapers testing the platform, but before we can take off we need more titles. It is a bit slow with newspapers. They are investing in their own individual platforms and might not see our platform to be interesting.

As the quote illustrates, the content provider is hesitant to newspapers interest in their platform. The informant continues to describe under what conditions he thinks newspapers might take an interest:

When the newspaper publishers realize that the print market is declining that also creates an interest for new solutions. The attitude has been a bit like do not fix it until its broken. But now the curves for print is only pointing down so we experience a bit more interest from newspapers, especially in Germany and England which enables us to provide a

large range of newspaper brands. But in for example Sweden there is still not much interest. But I think they will come to us too when they suffer enough losses in print.

The informant suggests that newspapers think they have a strong relationship to their readers and that they do not need to provide their content on outside their own individual platforms. According to the platform provider, newspapers are overestimating their relation to their readers, and under estimating the value of their content and brand. The paradox is that readers like to get their news from a well known brand because it is trusted and expected to provide quality content. If there are alternatives the informant does not think readers are loyal, so if the quality is compromised the readers are not loyal. The only loyal customers today are, according to the informant, old local readers. Another thing that the informant regards to be over estimated by newspaper is journalism:

Newspapers regard themselves to be a journalistic product, and not to be an edited product. Journalism is raw material to me, it can be bought on a market of media content and provided by readers. What makes a newspaper really good is to select, edit and present a good news service. That is what I think is the core value and what newspapers are best at, but they do not recognize that themselves.

The informant expects newspapers to have a role in the digital media ecosystem in the future. However, big actors such as Google or Apple, and new actors we do not know yet, will set the frames for how the ecosystem develops. New roles emerge, such as the broker role of platform owners, and the technology will be the driver of development. The basic media consumer needs are the same, according to the platform provider. The informant regards ownership structures, and reluctance to new co-operations, to be a risk for incumbent newspaper organizations in taking a role in the ecosystem:

A risk for the newspaper publishers is that they are too focused on ownership and only co-operate within their respective ownership structures. They forget that what is good for the readers is good for the staff, and that that is in turn good for the owners. If they prioritize the interest of the owners, the circle turns.

To leverage value in a digital platform ecosystem the informant considers co-operation in several constellations to be the necessary. There are many different actors in the ecosystem that can add competencies, infrastructures, markets etc. that one actor cannot build up on its own. This is according to the platform provider difficult for newspapers:

When we negotiate with newspapers they always want to take the dominant role. They are stuck in value-chain thinking and are unable to see how they can interact with other actors on equal terms. That is my experience. That is why we do not spend so much time to seek up newspapers, even though we know we can provide added value besides being a market place. We can provide aggregated and detailed information about reading patterns that can inform how to develop their products. We can offer them a younger audience then they have. But they do not see that.

The informant continues with a reflection on identity:

The biggest enemy for newspaper in the digital platform ecosystem is their inability to let go of the old identity as a dominant actor that rules the game. The time for that is over, and those who will make it the informant believes to be those who can size down and see themselves as one actor among others and join many different constellations in the ecosystem.

4.2 The newspaper informants

The newspaper informants provided a similar picture of the situation as the platform provider. In the group discussions the informants reflected on the content providers statements and of their future role in digital platform ecosystems. One of the informants had experience from negotiating with the platform provider:

We were in a discussion with the platform provider. We could not accept the share of 70/30, we think it should be the other way around. Compared to the margin of a printed newspaper this is nothing.

The statement was met by a question from another informant:

But then you do not have any distribution cost so how can you compare like that? I think the main problem is that they are in control of the customer relation. They bill the customer and know who they are, not we. That is the value we lose.

On the question of how they have approached new digital initiatives and offers by other actors, one informant described the following:

To be honest, we have not looked around us to see if there are any interesting collaborations. We have all been fumbling around looking for the Holy Grail. Is this the new digital business model, is that the one. Then we said, we need many complementing business models to cope with the transition. But have we really explored the ecosystem, or are we just fumbling around trying to do what we have always done?"

Another informant continued to describe that newspapers have been rigid and weighed down by the traditional print business model. He described they have been unable to act with the flexibility needed in digital ecosystems:

We are really stuck in doing business the print way. We say we have digital strategies, and that we are digital publishing companies but look at us. How digital are we really? Not very.

The informants were asked to respond to the platform providers statement on structures of ownership. One informant responded as follows:

The statement about ownership is so true. But at least we are starting to co-operate within our different owner structures now, and we have accomplished some new partnerships with digital companies. Still, we cannot see that it matters in figures yet.

Another informant described that part of the problem was related to the ability to join networked business and revenue shares:

I think we are scared to death that someone else will make a profit of our content. That is why we are so bad at networked business. We talk about digital transition. I do not think that is the core. The core is the transition from "every man for himself" to networked collaboration. It is not about technology.

One of the informants responded to this statement as follows:

Yes, I agree. If we had seen this development in time we would have a joint platform for newspaper content with a smart business model and convenient for readers and advertisers. We are so self-absorbed that we have all thought we can do it on our own.

The newspaper informants, however, took another stance than the platform provider regarding the role of journalism in value creation. They agreed upon that journalism is the core value they have to bring in the digital platform ecosystem, and that they can produce and offer quality content not only to readers but also to other actors in the ecosystem. They discussed this as an unexplored opportunity. One reason for this situation was, according to the informants, that they have to be too focused on providing content under their own brand. Another reason was that newspapers have not found their role and identity in digital platform ecosystems. One informant stated:

One thing that has been ambivalent for us is how we should define ourselves in digital ecosystems. Who are we, and who are our competitors. Sometimes we talk about Google and Facebook as competitors; sometimes it is computer games or even the local grocery store. I think we have an identity crisis. We do not know who our friends and enemies are and that makes it challenging to collaborate with other actors.

Another major challenge for newspapers in digital platform ecosystems was, according to the informants, the difficulty of charging readers for content. There have been a number of initiatives for pay walls, charging for specific services, but the revenues do not yet compensate for the loss of print revenues. The informants discussed this subject, how it affects the business models for services in digital platform ecosystems. As one informant put it:

One thing is for sure; if we cannot finance our own platforms with advertising revenues, we will be forced to join other platforms. Readers will not pay 10 euros per month for one brand, if they can have 100 brands for the same price on another platform. We can only keep our individual platforms if content is free to readers. That is what I think anyway.

On the topic of what has been the inertia for newspapers digital initiatives the informants discussed the dominant business strategy. The dominant business strategy was described as taking control over the value chain. One informant described the dominant strategy as follows:

The dominant business strategy is to be in control of the value chain. We have not been able to let go of that attitude and I think that is how our owners still strategize. The more control, the better. The more exposure and collaboration, the higher risk. If we can make it on our own we do, if we cannot we might consider some partnership. It has opened up, but we are still governed by this mind-set. When we assess a new initiative we compare to the print business, so the new initiative almost always loses. What business digital model could be as strong as the print model was? That is utopia.

4.3 The analysts

The analysts reflected on the future of newspaper publishers in digital platform ecosystems. Independence was put forward as important and valuable for newspapers. According to one informant, the content marketing trend on the advertising market is an opportunity to newspaper publishers. That is, if they do it well and not compromise reader's trust.

The value newspapers can build in digital ecosystems is to maintain independence in relation to advertisers. The problem is that the readers do not seem to care, so it is tempting to stretch the borders. But if they go too far they will lose trust and then the reader will go somewhere else. There is a balance here, how close can you integrate with advertisers, and how do readers value the trustworthiness of newspapers?

One analyst argued that newspapers spread their content too casually in social media (e.g. Twitter). If newspapers have the ambition to build strong brands in digital platform ecosystems, his advice was to be selective with where to spread content.

We are still in the childhood of these media. There is a lot of trial and error going on. The problem is though that I do not think they know why or why not to do things. Spreading things fragments newspaper content and makes newspapers lose control over the publishing context.

An important aspect of managing the transition to digital platform ecosystems is to attract a younger audience. Newspapers have, however, not managed to attract a critical mass of young readers in digital platforms. According to one of the informants, this is one of the most important factors if newspapers want to be attractive partners in the ecosystem:

Young people seem to be attracted by news media that is more entertainment oriented. That is not new. BuzzFeed is sometimes pointed to as the one who has managed to crack the youth code. Newspapers copycat their approach. What they need to consider though is that youngsters today have unlimited access to content and then they choose the candy instead of the veggies (veggies being traditional journalistic content). This is a major challenge if newspapers want to be attractive partners in the ecosystem.

Two of the analysts emphasized that it is very difficult for newspapers to enter other platforms with their content. They said not to be sure it is a wise strategy, as one of the informants stated:

I am not sure they (newspapers) should. Look at the music industry; the artists do not seem to be happy with the digital music platform business, do they? I think first and foremost newspapers should put their effort into renewal, but they are very conservative. They say they are not, but look at them. How many innovative digital initiatives have started with newspapers?

One analyst stressed the importance of anchoring on local markets. To have a position, newspapers need to be the best and the first. Otherwise grass root initiatives will emerge in the ecosystem and challenge the traditional newspapers. For society, the informant thinks both have a place:

I do not think it is good if there is no place for newspapers. Producing unique and relevant content is key, but the question is how to do that and how to finance that.

Attracting a large enough audience to be attractive for advertisers is challenging in digital platforms:

I think newspapers now must come to terms with that the golden age is over. They will never, ever again have as large proportions of local readers as they used to have.

One problem with the newspaper transition to digital platform ecosystems put forward is that they are too late to bring in the competencies they need:

They should have done what they do now 5-6 years ago, that is build up digital competence for mobile, web TV, advertising analytics etc. The strong old business model has blindfolded them. And when they have put in efforts, they put all eggs in one basket such as the tablet initiatives. That was an expensive experience for many of them.

Newspapers difficulty to collaborate was also described by the analytics:

They have been so occupied with competing and quarrel with each other, and between owners, that they miss the train over and over again. They have not managed to see themselves as complementary actors and together try to solve the problems and become more attractive to advertisers and other actors. Instead they fight each other.

One conclusion made was that it seems messy at this time, but that in time quality journalism will find its forms and business models in digital platform ecosystems. According to the three analytics there is a need for this. They also suggested that newspapers would not be the same news media we have been used to see (in print). As one of them put it:

Newspapers will rather be an exclusive content service. This opens for new businesses such as selling journalistic services to other actors.

5 Interpretations and discussion

With this study we aim to shed light on the challenges for incumbent content providers to adapt to and leverage business opportunities in digital platform ecosystems. With this aim we have studied how newspapers as content providers are challenged by the generativity afforded in digital platforms. The paradoxical relationship following the generative attribute of digital platforms challenges newspapers, and in particular this study display the challenges related to identity crisis and balancing openness and control. In the following we interpret and discuss these challenges in light of the generative attribute of digital platforms.

5.1 Newspapers identity crisis in digital platform ecosystems

Newspapers have had an identity of being in a dominant position and in control of the value chain. Now they struggle to survive in the digital ecosystem; dethroned, pressed to operate in the paradoxical relationship between the digital innovation opportunities and the historically well working traditional business model, while revenues are falling. As this study shows, this has led to an identity crisis, uncertainty of their role in the digital ecosystem, and to difficulties in explaining their role to others. While other key ecosystems actors (e.g. platform providers) recognize and embrace their identity, newspapers desperately seek to re-shape their identity and core business. This uncertainty is exacer-

bated due to increasing economical pressure, incumbent behavior and inertia in meeting digital platform opportunities (Christensen et al., 2012). A majority of the informants in this study state that newspapers have not been able to innovate neither technically nor business wise in such a way that they have enabled open and networked business models. Two plausible reasons for this inertia that this study points at is the unwillingness to accept the business conditions in digital platform ecosystems, and a “the more control the better” mind-set. While this unfolds, other ecosystem actors point out that newspapers have content, qualities and competence that is attractive and can provide value in the ecosystem. Content providers, for example, have incentives to enroll newspapers (content) in their platforms, but have experienced difficulties to engage newspapers. These difficulties can be related to insufficient understanding of the ecosystems' state and newspapers incumbent behavior.

One class of platform providers do not produce content, or at least to a limited extent. These platform providers act as brokers of media content. The actors that have control over the platform, and thus control the frames for business, have a key position in the ecosystem (Ghazawneh and Henfridsson, 2015). Newspapers are challenged by this architecture since they have built up solitary platforms for their own content, as it has for example the music industry. The newspaper informants express that the attitude is becoming more humble and open, and recognize that the past dominant position is part of the past. In self-reflection they see that they have not had any digital strategies in earnest, and have not had any strategies to build relationships with for example digital platform providers. This is partially explained to be a result of the history of control and the informants expressed that they are reluctant to any other actor profiting on their content. As discussed in literature (see e.g. Eaton et al., 2011), control in digital platforms can be a strategy to create competitive barriers for rivals, but this control can also result in unexpected new competition. The analytics point out that spreading and sharing content too willingly can lead to loss of control and open for innovative utilizing of newspaper content beyond newspaper's control.

5.2 Openness and platform control in digital platform ecosystems

In digital platform ecosystems content providers are challenged by the balance between controlling and stimulating platform generativity (Förderer et al., 2014). On the one hand newspapers have full and tight control within their own platforms, on the other hand their platforms are integrated in the open architecture of the Internet, and thus the content, if published without walls, accessible to anyone. This has since the introduction of digital news content in the 90s been challenging, and newspapers are still today struggling with it. The newspapers in the study recognize that there is a need to open up for networked business and that actors, such as the platform provider, are of interest to them. They do however also show difficulty in accepting the terms of networked business in the ecosystem.

The resolution of the previously tight coupling between distribution, media and content (Yoo et al., 2010) challenges newspapers in controlling digital initiatives and innovation. In digital platform ecosystems, the control is distributed among many different actors in complex and heterogenous networks (Eaton et al., 2015). These actors have divergent interests and driving forces, and in this business environment the previously firm centric business model of newspapers does not work in the same way. As suggested in literature, these new actors are disrupting but they can also be an opportunity to establish new relationships and leverage the innovation capability afforded by generativity (Tilson et al., 2010; Yoo, 2013).

In this study, the newspaper informants express an insistent need to explore digital platform ecosystems and the need to innovate and reflect upon how to build a portfolio of business models in the ecosystem. This is reflected in some initiatives for partnerships and cooperation. The analytics confirm this picture and state that newspapers did not see the changes in the ecosystem early enough to build up strategic partnerships for networked business. Now they are facing the same situation regarding advertising platforms. According to the analytics in the study, the newspapers have been too conservative, which has hampered innovativeness. However, the analytics recognize the challenge of profiting on digital newspaper content with the new competition and changes in the ecosystem. The resolved

coupling provides flexibility and potential for new re-combinations, new forms for distribution, new business models etc. This is one of the major driving forces behind digital innovation (Yoo, 2013) that can contribute to renew newspaper media. This opportunity also creates challenges in the form of competition from big and established actors such as Google, and new or unexpected competition from for example digital start-ups or grass root initiatives.

6 Concluding remarks

In conclusion, newspapers are challenged in digital platforms ecosystems as a consequence of the architecture and generative attribute of digital platforms which provides a contrasting business environment compared to the long history of the controlled business of print newspapers. This study contributes to literature (see e.g. Ghazawneh and Henfridsson, 2015; Förderer et al., 2014) by illustrating the relationships and reasoning behind the strategizing underlying decisions taken by incumbent content providers such as newspapers. We can conclude that the fundamental challenges associated with the complexity of balancing the old and new platform based business in a constantly evolving ecosystem is related to relational as well as technical aspects of digital platforms. In this study we have contributed by showing the challenges associated to identity in the digital platform ecosystem, and the balance between control and openness.

The generativity of digital platforms affords business opportunities of completely new character to incumbent content providers. One practical implication of the findings is that the balance of control and openness is an essential part of strategizing digital platforms, as is the identity in digital ecosystems. To follow the logic of digital platforms incumbent content providers need to re-evaluate their identity and their role in relation to other actors, as well as to communicate their identity to be better understood by others.

Future studies could address several limitations in our work. First, this study builds on accounts in relation to one type of content providers, i.e. newspapers. The challenges brought by the generativity afforded in digital platforms can be compared with other content providing industries such as magazines, books and films. Second, in this study there is only one informant representing a platform provider. Complementing insights on the relationship between content providers and platform providers can be gained by studying the relationship with more emphasis from the platform provider viewpoint.

References

- Adner, R. (2006). "Match your innovation strategy to your innovation ecosystem." *Harvard Business Review*, 84 (4), 98-107.
- Avital, M., and Te'eni, D. (2009). "From Generative Fit to Generative Capacity: Exploring an Emerging Dimension of Information Systems Design and Task Performance." *Information Systems Journal*, 19 (4), 345-367.
- Boczkowski, P. J. (2004). *Digitizing the News: Innovation in Online Newspapers*. Cambridge, MA, MIT Press.
- Chesbrough H. W. (2006). "Open innovation: A New Paradigm for Understanding Industrial Innovation." in Chesbrough, H., Vanhaverbeke, W. and West J. (eds.), *Open Innovation: Researching a New Paradigm*. Oxford: Oxford University Press, 1-12.
- Christensen, C. M., and Davis, A. B. (2006). "The Future of Newspapers." *The Forbes* URL: http://www.forbes.com/home/leadership/2006/10/10/leadership/newspapers/medialead/innovationcx_cc_1011clayton.html
- Christensen, C. M., Skok, D., and Allworth, J. (2012). "Breaking News: Mastering the Art of Disruptive Innovation in Journalism." *Nieman Reports*, 66 (3).
- Conboy, M. and Steel, J. (2008). "The future of Newspapers." *Journalism Studies*, 9 (5), 650-661.
- Eaton, B., Elaluf-Calderwood, S., Sørensen, C. and Yoo, Y. (2011). "Dynamic Structures of Control and Generativity in Digital Ecosystem Service Innovation: The Cases of the Apple and Google

- Mobile App Stores.” LSE working paper series, Draft available at: <http://is2.lse.ac.uk/wp/pdf/WP183.PDF>
- Eaton, B. D., Elaluf-Calderwood, S., Sørensen, C, and Yoo, Y. (2015). “Distributed Tuning of Boundary Resources: The Case of Apple’s iOS Service System.” *MIS Quarterly*, 39 (1), 217-243.
- Förderer, J., Kude, T., Schütz, S., and Heinzl, A. (2014). “Control versus Generativity: A Complex Adaptive Systems Perspective on Platforms.” In *Proceedings of ICIS 2014*. Auckland, New Zealand.
- Ghazawneh, A. and Henfridsson, O. (2013). “Balancing Platform Control and External Contribution in Third-Party Development: The Boundary Resources Model.” *Information Systems Journal*, 23 (2), 173-192.
- Ghazawneh, A. and Henfridsson, O. (2015). “A paradigmatic analysis of digital application market-places.” *Journal of Information Technology*, 30 (3), 198-208.
- Ihlström Eriksson, C., Kalling, T., Åkesson, M. and Fredberg, T. (2008). “Business Models for m-services - exploring the e-newspaper case from a consumer view.” *Journal of Electronic Commerce in Organizations*, 6 (2), 29-57.
- Küng, L. (2008). *Strategic Management in the Media: Theory to Practice*. Sage Publications.
- Noricom statistics, URL: <http://www.nordicom.se/en/media-trends/media-statistics>
- O’Reilly III, C. A. and Tushman, M. L. (2004). “The Ambidextrous Organization.” *Harvard Business Review*, 82 (4), 74-81.
- Orlikowski, W. J. and Baroudi, J. J. (1991). “Studying Information Technology in Organizations: Research Approaches and Assumptions.” *Information Systems Research*, 2 (1), 1-28.
- Tilson, D., Lyytinen, K., and Sørensen, C. (2010). “Research commentary-digital infrastructures: the missing IS research agenda.” *Information Systems Research*, 21(4), 748-759.
- Picard, R. G. (2003). “Cash Cows or Entrecôte: Publishing Companies and Disruptive Technologies.” *Trends in Communication*. 11 (2), 127–136.
- Picard, R. G. (2006). “Capital Crisis in the Profitable Newspaper Industry.” *Nieman Reports*, 60 (4), 10-12.
- Stone, M., Nel, F., McBride Mensching, L. and Lin, E. (2012). “World Newsmedia Innovation Study 2012.” Chicago, IL: *The World Newsmedia Network*.
- Tiwana, A., Konsynski, B., and Bush, A. A. (2010). “Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics.” *Information Systems Research*, 21 (4), 675–687.
- Walsham, G. (1995). “Interpretive case studies in IS research: nature and method.” *European Journal of Information Systems*, 4, 74-81.
- Walsham, G. (2006). “Doing interpretive research.” *European Journal of Information Systems*, 15(3), 320-330.
- Yoo, Y. (2013). “The Tables Have Turned: How Can the Information Systems Field Contribute to Technology and Innovation Management Research?” *Journal of the Association for Information Systems*, 14(5), 227-236.
- Yoo, Y., Boland, R.J., Lyytinen, K., and Majchrzak, A. (2012). “Organizing for Innovation in the Digitized World.” *Organization Science*, 23 (5), 1398-1408.
- Yoo, Y., Henfridsson, O. and Lyytinen, K. (2010). “The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research.” *Information Systems Research*, 21 (4), 724-735.
- Zittrain, J. (2006). “The Generative Internet.” *Harvard Law Review*, 119, 1974-2040.
- Åkesson, M. (2009). “Digital Innovations in the Value Networks of Newspapers.” PhD thesis. University of Gothenburg.
- Åkesson, M. and Ihlström Eriksson, C. (2009). “Advertising Challenges in Ubiquitous Media Environments.” In: Pousttchi, K.; Wiedemann, D.G. (eds.) *Handbook of Research on Mobile Marketing Management*. Information Science Reference, Hershey, 77-93.

FROM UNDERSTANDING TO USE AND COMPETE: A TRANSLATIONAL PLATFORM FOR BUSINESS TRANSFORMATION

Completed Research

Viscusi, Gianluigi, École Polytechnique Fédérale de Lausanne (EPFL - CDM MTEI CSI),
Lausanne, Switzerland, gianluigi.viscusi@epfl.ch

Tucci, Christopher, École Polytechnique Fédérale de Lausanne (EPFL - CDM MTEI CSI),
Lausanne, Switzerland, christopher.tucci@epfl.ch

Kokkinakos, Panagiotis, National Technical University of Athens, Zografou, Greece,
pkokkinakos@epu.ntua.gr

Alvertis, Iosif, National Technical University of Athens, Zografou, Greece,
alvertisjo@epu.ntua.gr

Koussouris, Sotiris, National Technical University of Athens, Zografou, Greece,
skous@epu.ntua.gr

Fenareti Lampathaki National Technical University of Athens, Zografou, Greece,
flamp@epu.ntua.gr

Abstract

This article discusses a translational cycle and a translational platform which have been designed in the context of the FutureEnterprise project, a European Commission funded support action. One of the main strategic axes of the FutureEnterprise project is related to a specific focus on translational research activities, aiming to bridge academic and industrial research with Internet-based entrepreneurship and digital business innovation. The term ‘translational research’ appeared in Pubmed illustrates, for the first time around 1993 to identify the “translational gaps”, hindering the transformation of discoveries in the life sciences into improvements having societal profit from basic research. As for the management research, translational issues have been pointed out as relevant and critical factors within Academy of Management (AOM) research community, identifying two types of translational challenges for an effective impact of management research on practice: a “lost in translation” (fail to find the right way to transfer research results in the practitioners language, understanding, and needs) and “lost before translation” (fail to identify an appropriate and systematic translation process as the one leading from “bench to bedside” in life sciences). The contribution presented in this article aims to face the challenges of ‘translational research’ in the context of technology management and innovation from a design science stance, thus identifying key constructs further developed through a translational platform which represents the resulting IT artifact (existing MOOC) from a “tool view”.

Keywords: Translational research, MOOC, Future Enterprise, Digital Platforms, Design Science, Digital Enterprise, Entrepreneurship.

1 Introduction

In this article we discuss a translational cycle and a translational platform which have been designed in the context of the FutureEnterprise project, a European Commission funded support action. One of the main strategic axes of the FutureEnterprise project is related to a specific focus on translational activities, aiming to bridge academic and industrial research with Internet-based entrepreneurship and digital business innovation. The contribution presented in this article aims to face the challenges of ‘*translational research*’ in the context of technology management and innovation from a design science stance (Hevner, March, Park, & Ram, 2004; March & Smith, 1995; Simon, 1996), thus identifying key constructs further developed through a translational platform which represents the resulting IT artifact from a “tool view” (Orlikowski & Iacono, 2001; Weber, 2003).

The article is structured as follows. First, background and motivations for the research are discussed along with an analysis of the MOOCs on web and digital business skills that are currently offered. Then, the research method is introduced before the FutureEnterprise translational cycle and platform are presented. Finally, conclusive remarks and future work discussion end the article.

2 Background and motivations

The debate on the role and difference among types of research has been at the centre of policy debate since the end of second World War, when the ideological stance on the pure value of basic science has been coupled with the idea of being it a necessary condition for technological advance as applied science, resulting among others from the reception of the 1945 Vannevar Bush report “Science-The Endless Frontier; A Report to the President on a Program for Postwar Scientific Research” (see the discussion in Stokes, 1997). This position has been questioned and challenged among others by Stokes (1997) with the identification of a quadrant model of scientific research along two main questions concerning research inspiration: is it a quest for fundamental understanding? Are there considerations of use? Thus what has been there called the Pasteur’s Quadrant was made up by pure basic research, pure applied research, use-inspired basic research (guided by technological needs), pointing out the role of the latter as well for creation of new disciplines and inspiration for basic research.

Yet, besides the investigation on the roles and types of research an increasing interest has risen by both academics and managers as well as entrepreneurs around *technology transfer (TT)*, whose definition strictly depends on the disciplines considering it (Bozeman, 2000; Reisman & Zhao, 1991; Zhao & Reisman, 1992): for example, management literature associates TT to competitive advantage and considers it as the transfer of specialized know-how, either patented or nonpatented, encompassing a sustained relationship between actors, e.g., enterprises; whereas social science and anthropology research link TT to diffusion of innovation or cultural change (Zhao & Reisman, 1992). As for innovation, in general, we can define it as the implementation of a new idea that leads to a change in practice in order to create some kind of value (Link & Siegel, 2007), referring either to the output or the process itself. Consequently, research on TT has produced evaluation models to understand the effects of this multifaceted phenomenon such as, e.g., the *Contingent Effectiveness Model* (Bozeman, Rimes, & Youtie, 2015; Bozeman, 2000). As pointed out by Bozeman (2000) the effectiveness of technology transfer should consider diverse determinants, including characteristics of the *transfer object, agent* (e.g., the nature of the institution, its history, and culture), *media* (e.g., patents, copyright, personnel exchange, spin-off, etc.), *recipient* (e.g., physical technology, scientific knowledge, design process) as well as the conditions related to the *demand environment*. However, the technology transfer effectiveness can have different meanings (e.g., market or political impacts) as actually has the definition of technology itself (Bozeman, 2000), spanning from “configurations” of processes and products/services to common sense view of technology as “tool” (Orlikowski & Iacono, 2001).

Nevertheless, however the efforts for linking TT to societal benefits and Public Value (see, e.g., Bozeman, Rimes, & Youtie 2015), the as pointed out by Bozeman et al. (2015) most of the studies on TT have focused on the United States (US) laboratories and research centres, transfer from university settings, multi-organizational research centres, consortia (thus, mainly international relations and owner technologies), or non-linear technology transfer (considering, e.g., relationship between university, industry, an government actors; open innovation; open source).

Taking the above issues into account, we now consider translational research as a potential complement to TT perspective, focusing specifically on the above mentioned types of research and gaps among them. The term '*translational research*' appeared in Pubmed illustrates, for the first time around 1993 (van der Laan & Boenink, 2012) to identify the "translational gaps", hindering the transformation of discoveries in the life sciences into improvements having societal profit from basic research. Among others, it is worth mentioning the following definition of translational research by Rubio et al. (2010) :

Translational research fosters the multidirectional integration of basic research, patient-oriented research, and population-based research, with the long-term aim of improving the health of the public. T1 research expedites the movement between basic research and patient-oriented research that leads to new or improved scientific understanding or standards of care. T2 research facilitates the movement between patient-oriented research and population-based research that leads to better patient outcomes, the implementation of best practices, and improved health status in communities. T3 research promotes interaction between laboratory-based research and population-based research to stimulate a robust scientific understanding of human health and disease.

As for the management research, translational issues have been pointed out as relevant and critical factors by Shapiro, Kirkman, & Courtney (2007) within Academy of Management (AOM) research community, pointing out two types of translational challenges for an effective impact of management research on practice: a "*lost in translation*" (fail to find the right way to transfer research results in the practitioners language, understanding, and needs) and "*lost before translation*" (fail to identify an appropriate and systematic translation process as the one leading from "bench to bedside" in life sciences).

Taking the above issues into account, we argue that among the key factors for translational processes there is education and, particularly today, instruments for higher global distance education such as Massive Open Online courses (MOOCs) are worth considering for promoting digital business orientation and web entrepreneurship skills (Alvertis et al., 2015). In what follows we discuss figures and insights on this phenomenon.

2.1 Massive Open Online courses (MOOCs)

What is the current offering for training and education focused on web and digital business skills and capabilities? That's the question guiding the survey the authors conducted in the late spring 2015. The research and analyses have been carried out along five macro-topics, suitable to provide the learning building blocks for a digital entrepreneurship curriculum:

- Business Strategy
- Entrepreneurship
- Innovation
- Management of Technology and Information Systems
- Social Studies of Technology

The choice aims to allow the identification of the "education seeds suitable" to provide the necessary skills and capabilities for entrepreneurs facing the systemic challenges posited at different levels by each domain. Also, the selection has been based on dimensions such as

- *Openness* (“Type of access” and “Fee for certificate” attributes) and
- *Accessibility* (“Main Language”, “Other Language”, and “Subtitles” attributes).

As for the openness dimension, the selection starts by considering Massive Open Online Courses (MOOCs), that are online courses for an unlimited participation and open access via the web, such as videos, readings, and problem sets as well as interactive user forums. MOOCs provide the chance for certification by international, often high ranked, academic institutions.

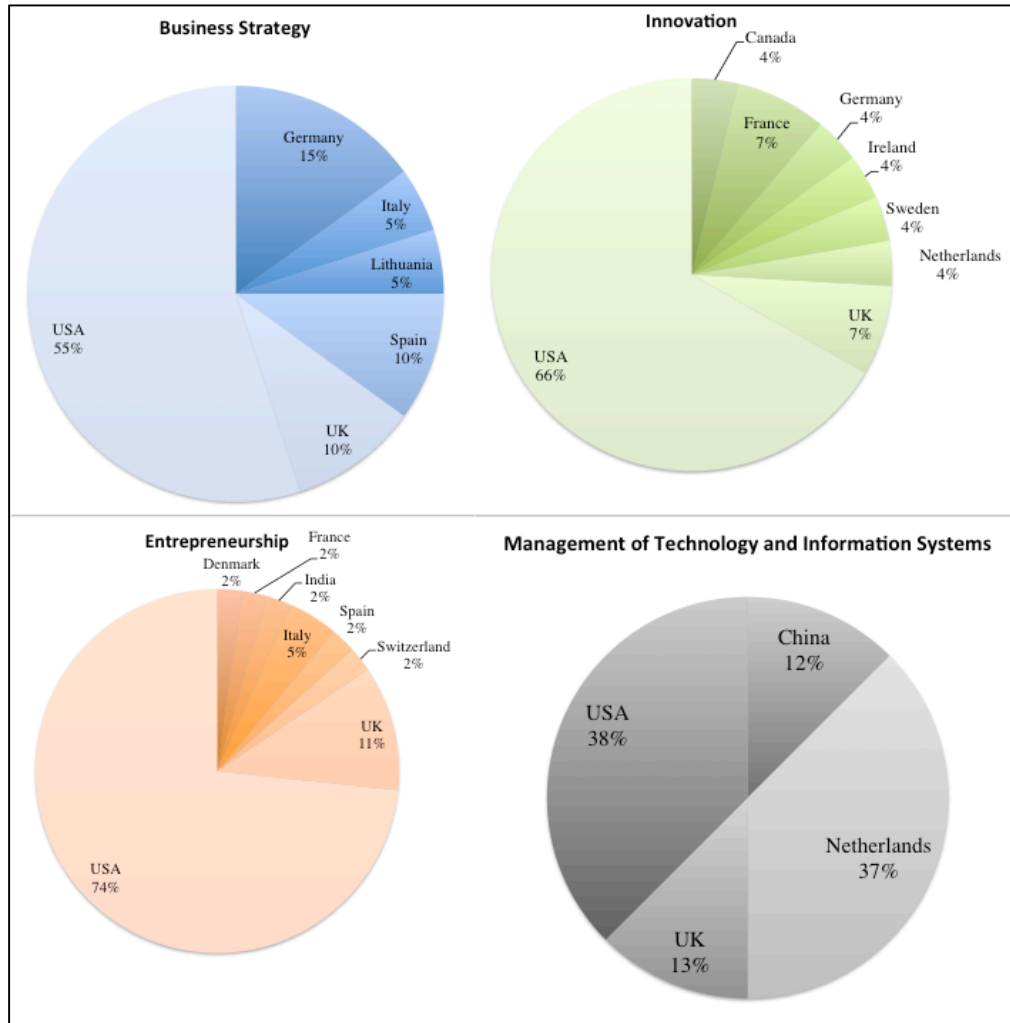


Figure 1. Offering per country per macro topic (%).

Accordingly, MOOCs represent an opportunity for entrepreneurs and enterprises willing to acquire skills and expertise they lack in a flexible and, in most cases, cost effective way. Besides the (obvious and expected) utilisation of popular search engines (such as Google), the authors thoroughly searched in relevant and/or dedicated portals and repositories for relevant MOOCs and courses. The list of the aforementioned sources can be found below:

- OpenEducationEuropa (www.openeducationeuropa.eu/en/find/moocs)
- KhanAcademy (<https://www.khanacademy.org/>)
- OpenupED (<http://www.openuped.eu/courses>)
- Udemy (<https://www.udemy.com/>)
- Udacity (<https://www.udacity.com/>)
- P2PU (<https://p2pu.org/en/>)

- Coursera (<https://www.coursera.org/>)
- edX (<https://www.edx.org/>)

The research resulted in the documentation of one hundred (100) courses (see <http://futureenterprise.eu/education-seeds>) leading to interesting insights presented in what follows (see also Figure 1). Among the considered macro topics, the most popular is “Entrepreneurship” by far, appearing in almost half of the cases (45 out of 100 cases). “Innovation”, “Business Strategy” and “Management of Information Technology” appeared 27, 20 and 8 times respectively. Considering the offering per country per macro topic (%), it is worth noting that UK is the only country (apart from the USA) covering all the macro topics. As for other European countries that cover more than one macro topic, the different focus of Germany (“Business strategy” and “Innovation”), Italy (“Business strategy” and “Entrepreneurship”), France (“Entrepreneurship” and “Innovation”) and the Netherlands (“Innovation” and “Management of Technology and Information Systems”) are worth noting.

2.1.1 Structure

Many of the identified courses are closed, as they were a one-off endeavour. There is a small number of cases that repeat after a substantial period of time. It is interesting to note that a large number of cases do not have specific dates; they are open and stakeholders are welcome to take the course whenever it is suitable for them. There is a case, though, of a relevant MOOC on Digital Enterprise by the Digital Business Academy that is restricted from a geographic perspective as it is available to UK residents only. Another important insight is that, although relevant courses exist for more than a decade, the majority of the reported cases took (or takes) place from 2014 onwards.

Regarding their structure, for most of the reported cases, a detailed overview of the course’s structure is provided to the prospective student, including the titles of both the sections and the subsections. In most of the cases, the courses are structured in four (4) to eight (8) sections, usually recommended to be followed on a weekly basis. It is also worth reporting that almost all courses have an introductory session as the first one, or a session providing an overview of the upcoming ones.

2.1.2 Providers

As for the organisations delivering the courses, universities seem to be the most active type of organisation by far; a result that was, however, expected as such initiatives totally fall in their interests and everyday activities. There was also a significant number of cases that were offered by individuals, without the support of a specific organisation. These individuals appeared to be of various different backgrounds: entrepreneurs, former professors, innovation managers, former senior executives etc. With regard to the gender issues, it is worth noting that in our sample, only 4 out 100 courses target specifically women (namely: “Make It Happen: Empowering Women for Success”, “Organizing Your Effort For The Ladypreneur”, “Positioning For Profit For The Ladypreneur”, “Women in Leadership: Inspiring Positive Change”), and mainly from institutions based in the USA.

In the cases where the courses were distributed by universities and organisations and not by individuals, there were rare occasions where the name of the instructor was not given. On such cases, it was considered adequate to indicate the name of the organisation or university by which the instructor will be drawn. Along these lines, it is important to note that organisations and individuals from the USA are particularly active in the domain, whereas, in the EU, universities play the leading role in the distribution of these courses. Considering the xMOOC market. Notwithstanding competitors as edX, FutureLearn, Udacity and Udemy, Coursera, founded in 2012 by computer science professors Andrew Ng and Daphne Koller from Stanford University, is actually the most diffused and used platform with 13,254,433 users from 190 countries enrolled and offered more than 1041 courses from 119 institutions (<https://www.coursera.org/> - last accessed on May 29th, 2015).

2.1.3 Audience

Almost all of the reported MOOCs and online courses targeted a very broad spectrum of stakeholders, including: Entrepreneurs, (Innovation) managers, Business analysts, Students, SMEs, Enterprises, and IT literate employees. Only single cases focus on specific audiences such as R&D departments of enterprises, young and/or female entrepreneurs; an interesting case is the MOOC that the Digi-fem project women is developing, shown in Table 1 with other European Union (EU) funded projects that develop MOOCs focused on web and digital business skills and capabilities.

| Name | Description |
|---|--|
| <i>Digi-fem</i> (http://digifem.eu/project-description/) | The Digi-fem project is going to develop a MOOC, targeting young women (from 18 to 38 years), NEETs, “with fewer opportunities”, with or without entrepreneurial activity, willing to be integrated in the labor market. |
| <i>Digistart</i> (http://digistart.unidemi.com) | The Digistart project is developing a collection of MOOCs aimed at motivating young entrepreneurs. A special section will focus on encouraging women to become online entrepreneurs by harnessing the use of the internet for expanding small home-based businesses into global enterprises. |
| <i>Flite</i> (http://flite-proj.cenfim.pt) | The Flite project is developing a course on “Entrepreneurship and Innovation” (currently at Pilot 2), allowing to learn about entrepreneurship in 50 hours over an 8 week period. Working in teams, the course participants should produce a business plan to form a startup company using the Business Model Canvas (Osterwalder & Pigneur, 2010) |
| <i>Heinnovate</i> (https://heinnovate.eu) | The project does not develop a MOOC per se, but provides self-assessment tool for Entrepreneurial Higher Education Institutions. |
| <i>InvestHorizon</i> (http://www.investhorizon.eu/online_courses.aspx) | InvestHorizon provides on-line courses as podcasts on launching the campaign - setting up the company structure, managing investor, managing and closing a financing round. |
| <i>LEAD - e-Leadership Skills</i> (http://www.eskills-lead.eu/home/) | LEAD is going to develop training materials focused on creating value from ICT such as cloud computing, mobile, social technology and Big Data. The programs will address key SME e-leadership skills gaps and demonstrate effective teaching of these. This especially includes competences required to initiate and guide innovation through ICT in an SME or start-up context. |
| <i>Lean LaunchPad Pilot</i> (http://www.europeanlaunchpad.com/) | The project provides entrepreneurship courses for the following areas : <ul style="list-style-type: none"> - <i>Open Disruptive Innovation(ODI)</i>: 6 weeks on-line training and dedicated on-site classes as follow-up; - <i>eHealth and Active&Healthy Ageing</i>: 10 weeks on-line training and dedicated on-site classes as follow-up. - <i>Future and Emerging Technologies</i>: 10 weeks on-line training and dedicated on-site classes as follow-up. It focuses on educating entrepreneurs on lean management in Europe. |
| <i>VET4STARTUP</i> (http://www.vet4startup.eu) | The VET4STARTUP is developing a MOOC (Massive Open Online Course) for start-uppers, with animated videos in 3 modules about the mind-set and skills needed to design an enterprise. |

Table 1. Relevant EU funded projects that develop MOOCs

Moreover, most of the MOOCs are offered in English and without subtitles. The second most popular language is Spanish, while also localised courses (e.g. in Ukrainian, Chinese, French, Portuguese) have been also identified and reported.

Finally, as the financial cost of the reported MOOCs and courses is concerned, most of them are free of charge. The same stood for the certificate they offered; however, it is interesting to note that there are courses that did not deliver any kind of certification to those completing the course. Contrary to that, all courses are accompanied by a specific price offered a certification (included in the price).

As a general conclusion, business and entrepreneurial activity concerns the vast majority of people and that could not be missing from the increasing presence of MOOC's on web, where there is abundant offer of courses around the fields of entrepreneurship and business. Furthermore, there is lack of a MOOC which integrates, in a single course, the key issues related to the three core macro topics for entrepreneurs in order to understand the trends impacting on business (Business Strategy), what are business models and how they can enforce innovation, the opportunities of open innovation (Innovation), together to value creation issue (Entrepreneurship). A closer offering could be one similar to what the Lean LaunchPad Pilot project provides, however not structured as a single MOOC, focusing on lean management and requiring still the attendance of diverse dedicated even if connected programs; thus requiring an engagement not always accessible in terms of availability of time and resources by the business learners.

3 Method

The research presented in this article adopts Design Science perspective (Hevner et al., 2004). Accordingly, the activities typically making up the steps of a Design Science research *building, evaluating, theorizing on and justifying* artifacts. The work presented in this paper concerns the identification of translation model with constructs (key factors and gaps) and a cycle representing the sequence related to their performativity (*building*). Thus, the article discusses the early stage of a design science research.

4 The translational cycle and platform

Taking the above issues into account, the FutureEnterprise consortium has designed a *Translational Cycle* which enabling “use-inspired basic research” (van der Laan & Boenink, 2012). The cycle is made up of three steps (Learn, Network and Compete) aiming to enforce solutions to translation gaps as follows (see Figure 2).

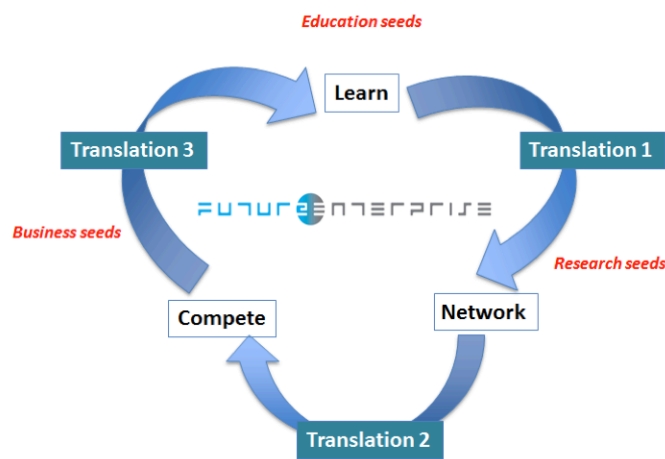


Figure 2. The Translational Cycle

Web or Digital Entrepreneurs (WDEs) are supposed to use *education seeds* as well as the *FutureEnterprise MOOC* to improve, extend, or consolidate their capabilities and skills in the *Learn*

step. The latter aims to provide a better understanding of what is actually available from research projects in the digital technology innovation domain (*Translation 1*). Thus, in the *Network* step, the WDEs can identify and contact suitable partners (*Research seeds*), fitting the opportunities identified for their entrepreneurial projects. This implies a common understanding of the resources, goals, and perspectives of the potential partners willing to design common new entrepreneurial projects (*Translation 2*). It is worth noting that between the considered translation steps a two-way interaction and feedbacks can happen, in order to align knowledge results from projects to the available academic education seeds.

Subsequently, the Web or Digital Entrepreneurs can either compete together with common products/services or participate to ideas competitions or crowdfunding initiatives promoted by international institutions/organisations such as, e.g., Kickstarter¹ or Innocentive (*Compete* steps). Both new common products/services and new ideas imply the design, identification, and adoption of innovations to business models actually available on the market and by the participants (*business seeds*). Also in this case a two-way interaction and feedbacks can happen (*Translation 3*). The flow adopted within the cycle encompasses what in translational research is mentioned as “backwards translation”, that is the results produced in a specific stage of research has to influence and be “feedback” to earlier phases of the cycle. It is worth noting that the above mentioned steps have a corresponding implementation as the core components of a translation platform that represent the IT artefact resulting from the design science research. Specifically the platform is implemented through the *FutureEnterprise Lab* discussed in what follows and shown in Figure 3.

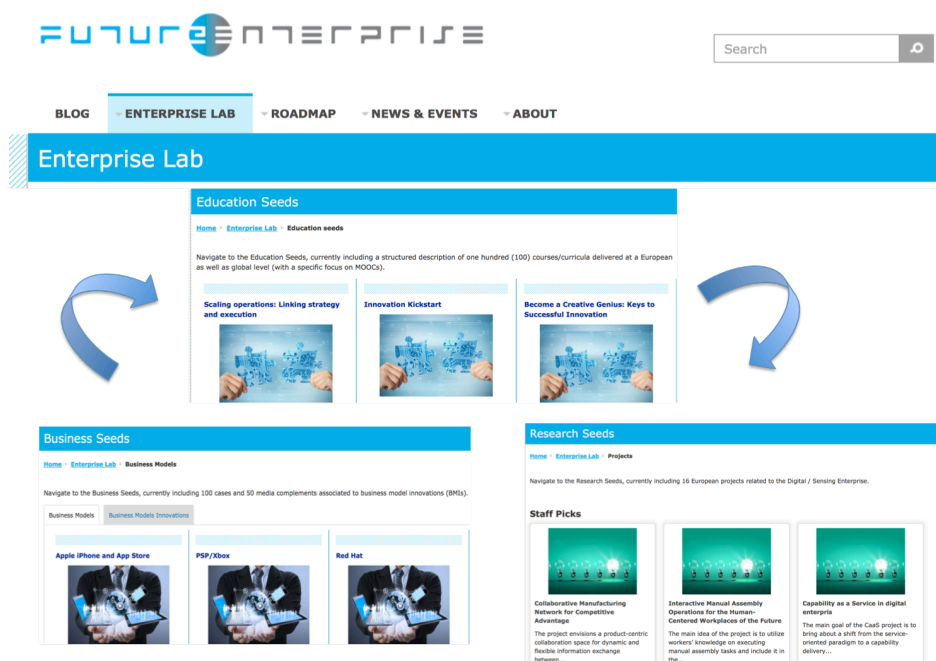


Figure 3. The Translational Platform

The translational cycle, as well as its associated steps and seeds, provide a framework for learning from best practices as they continuously emerge from research. Taking these issues into account, adopting, for example, a structuration theory perspective (Giddens, 1984) as theoretical lens, the translational cycle together with the FutureEnterprise Lab translation platform may be considered as the *rules* and *resources* (a *structure* in the structuration theory perspective), medium and outcome of

¹ <https://www.kickstarter.com/>

business innovation practices, mutually enacting enterprises and entrepreneurs (agents in the structuration theory perspective) diverse innovation systems.

5 Conclusion and future work

This article has discussed a translational cycle and translation platform which have been designed in the context of the FutureEnterprise project, a European Commission funded support action. The contribution is framed under a design science perspective and method, aiming to face the challenges of ‘*translational research*’ in the context of technology management and innovation. The translational cycle and platform they aim to provide a ground for inspiration of innovative new business ideas based on defined research outcomes, the latter considered a potentially effective mechanism to attract the long-lasting interest of potential young entrepreneurs and SMEs, as well as for the business digital transformation of large enterprises. It is worth noting that for the facilitation and homogenization of the laws and regulations among the various EU countries regarding cross-border entrepreneurship, the FutureEnterprise seeds may provide a better knowledge of the requirements for diverse business model innovations (Massa & Tucci, 2014) made available, e.g., by the digitalization of business (Tilson, Lyytinen, & Sørensen, 2010; Yoo, Henfridsson, & Lyytinen, 2010; Yoo, 2013), thus choosing the better business solution fitting the actual legal and juridical environment.

However these potential benefits, the limitations of the presented research concern the evaluation by the target users, which will be developed through focus groups, surveys as well as crowd based assessment initiatives. Thus, future work will concern the completion of the remaining steps actually characterising a design science research, that are *evaluating*, *theorizing on* and *justifying* artifacts.

6 Acknowledgments

This work has been partly funded by the European Commission through the FP7 Project FutureEnterprise “Road mapping, Research Coordination and Policy activities supporting Future Internet-based Enterprise Innovation” (Grant Agreement No. 611948).

References

- Alvertis, I., Kokkinakos, P., Koussouris, S., Lampathaki, F., Psarras, J., Viscusi, G., & Tucci, C. (2015). Challenges Laying Ahead for Future Digital Enterprises: A Research Perspective. In A. Persson & J. Stirna (Eds.), *Advanced Information Systems Engineering Workshops SE - 20* (Vol. 215, pp. 195–206). Springer International Publishing. doi:10.1007/978-3-319-19243-7_20
- Bozeman, B. (2000). Technology transfer and public policy: a review of research and theory. *Research Policy*, 29(4-5), 627–655. doi:10.1016/S0048-7333(99)00093-1
- Bozeman, B., Rimes, H., & Youtie, J. (2015). The evolving state-of-the-art in technology transfer research: Revisiting the contingent effectiveness model. *Research Policy*, 44(1), 34–49. doi:http://dx.doi.org/10.1016/j.respol.2014.06.008
- Giddens, A. (1984). *The Constitution of Society - Outline of the Theory of Structuration*. Berkeley and Los Angeles: University of California Press.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75–105.
- Link, A. N., & Siegel, D. S. (2007). *Innovation, Entrepreneurship, and Technological Change*. Oxford University Press.
- March, S., & Smith, G. (1995). Design and Natural Science Research on Information Technology. *Decision Support Systems*, (15), 251 – 266.

- Massa, L., & Tucci, C. L. (2014). Business Model Innovation. In M. Dodgson, D. M. Gann, & N. Phillips (Eds.), *The Oxford Handbook of Innovation Management*. Oxford Univ. Press. doi:0.1093/oxfordhb/9780199694945.013.002
- Orlikowski, W. J., & Iacono, C. S. (2001). Desperately seeking the “IT” in IT research: A call to theorizing the IT artifact. *Information Systems the State of the Field*, 34, 121–134.
- Osterwalder, A., & Pigneur, Y. (2010). *Business Model Generation. A Handbook for Visionaries, Game Changers, and Challengers*. John Wiley & Sons.
- Reisman, A., & Zhao, L. (1991). A Taxonomy of Technology Transfer Transaction Types. *Technology Transfer*, 16(2), 38–42.
- Rubio, D. M., Schoenbaum, E. E., Lee, L. S., Schteingart, D. E., Marantz, P. R., Anderson, K. E., ... Esposito, K. (2010). Defining translational research: implications for training. *Academic Medicine: Journal of the Association of American Medical Colleges*, 85(3), 470–475. doi:10.1097/ACM.0b013e3181ccd618
- Shapiro, D. L., Kirkman, B. L., & Courtney, H. G. (2007). Perceived Causes and Solutions of the Translation Problem in Management Research. *The Academy of Management Journal*, 50(2), 249–266.
- Simon, H. A. (1996). *The Sciences of the Artificial* (3rd ed.). Cambridge (Mass.): The MIT Press.
- Stokes, D. E. (1997). *The Pasteur’s Quadrant - Basic Science and Technological Innovation*. Washington D.C.: Brookings Institution Press.
- Tilson, D., Lyytinen, K., & Sørensen, C. (2010). Digital Infrastructures: The Missing IS Research Agenda. *Information Systems Research*, 21(4), 748–759.
- Van der Laan, A. L., & Boenink, M. (2012). Beyond Bench and Bedside: Disentangling the Concept of Translational Research. *Health Care Anal.* doi:10.1007/s10728-012-0236-x
- Weber, R. (2003). Still Desperately Seeking the IT Artifact. *MIS Quarterly*, 27(2), iii–xi.
- Yoo, Y. (2013). The Tables Have Turned: How Can the Information Systems Field Contribute to Technology and Innovation Management Research?. *Journal of the Association for Information Systems*, 14, 227–236.
- Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). The new organizing logic of digital innovation: An agenda for information systems research. *Information Systems Research*, 21(4), 724–735. doi:10.1287/isre.1100.0322
- Zhao, L., & Reisman, A. (1992). Toward Meta Research on Technology Transfer. *IEEE Transactions on Management Engineering*, 39(1), 13–21.

INFORMATION SYSTEMS IN A CHANGING ECONOMY AND SOCIETY

Proceedings of the 9th Mediterranean Conference on
Information Systems (MCIS 2015)
Samos, Greece, October 3-5, 2015

Editors

Spyros Kokolakis
University of the Aegean
Samos
Greece

Maria Karyda
University of the Aegean
Samos
Greece

Euripidis Loukis
University of the Aegean
Samos
Greece

Yannis Charalabidis
University of the Aegean
Samos
Greece

ISBN (ebook): 978-618-82208-6-7



UNIVERSITY OF THE AEGEAN