ICE-B 2010

International Conference on e-Business

Proceedings

Athens, Greece · 26 - 28 July, 2010

ICE-B 2010

Proceedings of the
International Conference on e-Business

Athens, Greece

July 26 - 28, 2010

Edited by David Marca, Boris Shishkov and Marten Van Sinderen

http://www.ice-b.icete.org

ice-b.secretariat@insticc.org

BRIEF CONTENTS

INVITED SPEAKERS

Ian F. Akyildiz

Georgia Institute of Technology

U.S.A.

Petar M. Djuric

Stony Brook University

U.S.A.

Stamatios Kartalopoulos

University of Oklahoma

U.S.A.

Nikolaos Bourbakis

Wright State University

U.S.A.

ORGANIZING AND STEERING COMMITTEES

CONFERENCE CO-CHAIRS

Mohammad S. Obaidat, Monmouth University, U.S.A.

George Tsihrintzis, University of Piraeus, Greece

PROGRAM CO-CHAIRS

David Marca, University of Phoenix, U.S.A.

Boris Shishkov, IICREST / Delft University of Technology, Bulgaria

Marten Van Sinderen, University of Twente / CTIT / IICREST, The Netherlands

PROCEEDINGS PRODUCTION

Patrícia Alves, INSTICC, Portugal

Helder Coelhas, INSTICC, Portugal

Vera Coelho, INSTICC, Portugal

Andreia Costa, INSTICC, Portugal

Patricia Duarte, INSTICC, Portugal

Bruno Encarnação, INSTICC, Portugal

Mauro Graça, INSTICC, Portugal

Liliana Medina, INSTICC, Portugal

Carla Mota, INSTICC, Portugal

Raquel Pedrosa, INSTICC, Portugal

Vitor Pedrosa, INSTICC, Portugal

Filipa Rosa, INSTICC, Portugal

José Varela, INSTICC, Portugal

CD-ROM PRODUCTION

Elton Mendes, INSTICC, Portugal

Pedro Varela, INSTICC, Portugal

GRAPHICS PRODUCTION AND WEBDESIGNER

Daniel Pereira, INSTICC, Portugal

SECRETARIAT

Mónica Saramago, INSTICC, Portugal

WEBMASTER

Sérgio Brissos, INSTICC, Portugal

PROGRAM COMMITTEE

Anteneh Ayanso, Brock University, Canada

Ladjel Belllatreche, Poitiers University, France

Morad Benyoucef, University of Ottawa, Canada

Indranil Bose, The University of Hong Kong, Hong Kong

Christos Bouras, University of Patras and RACTI, Greece

Stéphane Bressan, National University of Singapore, Singapore

Erik Buchmann, Universität Karlsruhe (TH), Germany

Rebecca Bulander, Pforzheim University of Applied Science, Germany

Wojciech Cellary, Poznan University of Economics, Poland

Dickson Chiu, Dickson Computer Systems, China

Soon Chun, City University of New York, U.S.A.

Jen-Yao Chung, IBM T. J. Watson Research Center, U.S.A.

Michele Colajanni, University of Modena and Reggio Emilia, Italy

Rafael Corchuelo, University of Sevilla, Spain

Peter Dolog, Aalborg University, Denmark

Khalil Drira, LAAS-CNRS, France

Yanqing Duan, University of Bedfordshire, U.K.

Erwin Fielt, Queensland University of Technology, Australia

Chiara Francalanci, Politecnico Di Milano, Italy

Flavius Frasincar, Erasmus University Rotterdam, The Netherlands

José María García, University of Sevilla, Spain

Paul Grefen, Eindhoven University of Technology, The Netherlands

Andreas Holzinger, Medical University Graz, Austria

Ela Hunt, ETH Zurich, Switzerland

Arun Iyengar, IBM Research, U.S.A.

Anton Lavrin, The Technical University of Kosice, Slovak Republic

Sebastian Link, Victoria University of Wellington, New Zealand

Liping Liu, University of Akron, U.S.A.

Jie Lu, University of Technology, Sydney, Australia

David Marca, University of Phoenix, U.S.A.

Tokuro Matsuo, Yamagata University, Japan

Gavin McArdle, University College Dublin, Ireland

Brian Mennecke, Iowa State University, U.S.A.

Adrian Mocan, SAP AG, Germany

Ali Reza Montazemi, Mcmaster University, Canada

Maurice Mulvenna, University of Ulster, U.K.

Dan O'Leary, University of Southern California, U.S.A.

Georgios Papamichail, New York College (Athens Campus), Greece

Krassie Petrova, Auckland University of Technology, New Zealand

Pak-Lok Poon, The Hong Kong Polytechnic University, China

Philippos Pouyioutas, University of Nicosia, Cyprus

Sofia Reino, Queen Margaret University, Edinburgh, U.K.

Ana Paula Rocha, LIACC-NIAD&R / FEUP, Portugal

Gustavo Rossi, University of La Plata, Argentina

Jarogniew Rykowski, The Poznan University of Economics, Poland

Quah Tong Seng, Nanyang Technological University, Singapore

Zhaohao Sun, University of Ballarat, Australia

Thompson Teo, National University of Singapore, Singapore

PROGRAM COMMITTEE (CONT.)

AUXILIARY REVIEWERS

SELECTED PAPERS BOOK

A number of selected papers presented at ICE-B 2010 will be published by Springer-Verlag in a CCIS Series book. This selection will be done by the Conference Co-chairs and Program Co-chairs, among the papers actually presented at the conference, based on a rigorous review by the ICETE 2010 Program Committee members.

FOREWORD

We warmly welcome you to ICE-B 2010 - the International Conference on E-business, which is held, this year, in Athens, Greece, hosted by the University of Piraeus.

ICE-B 2010 is integrated in the ICETE (International Conference on e-Business and Telecommunications) joint conference, which puts together six complementary conferences, namely DCNET, ICE-B, OPTICS, SECRYPT, SIGMAP and WINSYS, covering a broad range of related fields, including data communication networking, e-business, optical communication systems, security and cryptography, signal processing and multimedia applications, and wireless networks and information systems.

Sponsored by INSTICC (the Institute for Systems and Technologies of Information, Control and Communication), the conference is technically co-sponsored by IEEE Communications Society and IEEE Systems, Man and Cybernetics, and held in cooperation with IACR (the International Association for Cryptologic Research), ACM SIGSAC (Special Interest Group on Security, Audit and Control), ACM SIGMM (Special Interest Group on Multimedia), and IEICE (the Japanese Institute of Electronics, Information and Communication Engineers).

In the program for ICETE, we have included keynote lectures, papers, and posters to present the widest possible view on these technical areas. With its five tracks, we expect to appeal to a global audience of the engineers, scientists, business practitioners and policy experts, interested in R&D on Telecommunication Systems and Services:

- e-Business Applications
- Enterprise Engineering
- Mobility
- Business Collaboration & e-Services
- Technology Platforms

All tracks focus on research related to real world applications and rely on contributions not only from Academia, but also from industry, with different solutions for end-user applications and enabling technologies, in a diversity of communication environments. The proceedings demonstrate a number of new and innovative solutions for e-business and telecommunication, and demonstrate the vitality of these research areas.

ICETE has received 422 papers in total, with contributions from 46 different countries, in all continents, which demonstrates the success and global dimension of ICETE 2010. To evaluate each submission, a double blind paper evaluation method was used: each paper was reviewed by at least two experts from the International Program Committee, in a double-blind review process, and most papers had 3 reviews or more. The selection process followed strict criteria in all tracks so

only 48 papers were accepted and orally presented at ICETE as full papers (11.3% of submissions) and 74 as short papers (17.5% of submissions). Additionally, 60 papers were accepted for poster presentation. With these acceptance ratios, ICETE 2010 continues the tradition of previous ICETE conferences, based on high quality standards. The best papers of the conference will be invited to appear in a post-conference book that will be published by Springer-Verlag.

We would like to emphasize that ICETE 2010 includes several outstanding keynote lectures, which are relevant to today's lines of research and technical innovation. These talks are presented by distinguished researchers who are internationally renowned experts in one or more of the ICETE areas, and who help enhancing the overall quality of our Conference by their contributions.

A successful conference involves more than paper presentations; it is also a meeting place, where ideas about new research projects and other ventures are discussed and debated. Therefore, a social event including a conference diner/banquet has been planned for the evening of July 27 in order to promote this kind of social networking.

We would like to express our thanks to all colleagues involved in supporting this conference. First of all, we thank all authors including those whose papers were not included in the program. Next, we would like to thank all members of the program committee and reviewers, who provided an invaluable help with their expertise, dedication and time. We would also like to thank the invited speakers for their invaluable contribution, in the sharing their vision and knowledge. Naturally, a word of appreciation for the work of the secretariat and all other members of the organization, whose diligence in dealing with all organizational issues were essential and required a collaborative effort of a dedicated and highly capable team.

We hope that these proceedings may be a helpful reference for the future work for all those who need to address topics in the areas of e-business and telecommunications.

Enjoy the program and your stay in Athens.

David Marca
University of Phoenix, U.S.A.

Boris Shishkov
IICREST / Delft University of Technology, Bulgaria

Marten Van Sinderen
University of Twente / CTIT / IICREST, The Netherlands

CONTENTS

POSTERS

ENTERPRISE ENGINEERING

FULL PAPER

TECHNOLOGY PLATFORMS

SHORT PAPERS

POSTER

INVITED SPEAKERS

KEYNOTE SPEAKERS

NANONETWORKS
A New Frontier in Communications

Ian F. Akyildiz

Georgia Institute of Technology, U.S.A.

Abstract: Nanotechnology is enabling the development of devices in a scale ranging from one to a few one hundred nanometers. Nanonetworks, i.e., the interconnection of nano-scale devices, are expected to expand the capabilities of single nano-machines by allowing them to cooperate and share information. Traditional communication technologies are not directly suitable for nanonetworks mainly due to the size and power consumption of existing transmitters, receivers and additional processing components. All these define a new communication paradigm that demands novel solutions such as nano-transceivers, channel models for the nano-scale, and protocols and architectures for nanonetworks. In this talk, first the state-of-the-art in nano-machines, including architectural aspects, expected features of future nano-machines, and current developments are presented for a better understanding of the nanonetwork scenarios. Moreover, nanonetworks features and components are explained and compared with traditional communication networks. Novel nano-antennas based on nano-materials as well as the terahertz band are investigated for electromagnetic communication in nanonetworks. Furthermore, molecular communication mechanisms are presented for short-range networking based on ion signaling and molecular motors, for medium-range networking based on flagellated bacteria and nanorods, as well as for long-range networking based on pheromones and capillaries. Finally, open research challenges such as the development of network components, molecular communication theory, and new architectures and protocols, which need to be solved in order to pave the way for the development and deployment of nanonetworks within the next couple of decades are presented.

BRIEF BIOGRAPHY

Ian F. Akyildiz received his BS, MS, and PhD degrees in Computer Engineering from the University of Erlangen-Nuremberg, Germany, in 1978, 1981 and 1984, respectively. Currently, he is the Ken Byers Distinguished Chair Professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Director of Broadband Wireless Networking Laboratory and Chair of the Telecommunication Group at Georgia Tech.

Dr. Akyildiz is also an Honorary Professor with the School of Electrical Engineering at the Universitat Politècnica de Catalunya, Barcelona, Spain, since June 2008. Also since March 2009, he is an Honorary Professor with the Department of Electrical, Electronic and Computer Engineering at the University of Pretoria, South Africa.

He is the Editor-in-Chief of Computer Networks (Elsevier) Journal, the founding Editor-in-Chief of the Ad Hoc Networks Journal (Elsevier) in 2003, the founding Editor-in-Chief of the Physical Communication (PHYCOM) Journal (Elsevier) in 2008, and the founding Editor-in-Chief of Nano Communication Networks (NANO-COMNET) Journal (Elsevier) in 2010. Dr. Akyildiz serves on the advisory boards of several research centers, journals, conferences and publication companies.

Dr. Akyildiz is an IEEE Fellow (1996) and an ACM Fellow (1997). He received numerous awards from IEEE and ACM. His current research interests are in Nanonetworks, Cognitive Radio Networks, and Wireless Sensor Networks.

FROM NATURE TO METHODS AND BACK TO NATURE

Petar M. Djuric

Stony Brook University, U.S.A.

Abstract: A fundamental challenge in today's arena of complex systems is the design and development of accurate and robust signal processing methods. These methods should be capable to adapt quickly to unexpected changes in the data and operate under minimal model assumptions. Systems in Nature also do signal processing and often do it optimally. Therefore, it makes much sense to understand what Nature does and try to mimic it and do even better. In return, the results of better signal processing methods may lead to new advancements in science and technology and in understanding Nature. In this presentation methods for signal processing that borrow concepts and principles found in Nature are addressed including ant optimization, swarm intelligence and genetic algorithms. However, the emphasis of the presentation is on Monte Carlo-based methods, and in particular, methods related to particle filtering, cost-reference particle filtering, and population Monte Carlo. In the past decade and a half, Monte Carlo-based methods have gained considerable popularity in dealing with nonlinear and/or non-Gaussian systems. The three groups of methods share the feature that they explore spaces of unknowns using particles and weights (costs) assigned to the particles. In most versions of these methods, particles move independently and in accordance with the dynamics of the assumed model of the states. Interactions among particles only occur through the process of resampling rather than through local interactions as is common in physical and biological systems. Such interactions can improve the performance of the methods and can allow for coping with more challenging problems with better efficiency and accuracy. We show how we apply these methods to problems in engineering, economics, and biology.

BRIEF BIOGRAPHY

Petar M. Djurić received his B.S. and M.S. degrees in electrical engineering from the University of Belgrade, in 1981 and 1986, respectively, and his Ph.D. degree in electrical engineering from the University of Rhode Island, in 1990. From 1981 to 1986 he was a Research Associate with the Institute of Nuclear Sciences, Vinča, Belgrade. Since 1990 he has been with Stony Brook University, where he is Professor in the Department of Electrical and Computer Engineering. He works in the area of statistical signal processing, and his primary interests are in the theory of signal modeling, detection, and estimation and application of the theory to a wide range of disciplines. Prof. Djurić has been invited to lecture at universities in the United States and overseas and has served on numerous committees for the IEEE. During 2008-2009 he was Distinguished Lecturer of the IEEE Signal Processing Society. He was the Area Editor for Special Issues of the Signal Processing Magazine and Associate Editor of the IEEE Transactions on Signal Processing. He has also been on the Editorial Boards of many IEEE and EURASIP professional journals. In 2007, he received the IEEE Signal Processing Magazine Best Paper Award. Prof. Djurić is a Fellow of IEEE.

CHAOTIC QUANTUM CRYPTOGRAPHY
The Ultimate for Network Security

Stamatios Kartalopoulos
University of Oklahoma, U.S.A.

Abstract: As the sophistication of intruders' increases, so does the incidents of information integrity breaches and network attacks. In response, very complex cryptographic processes have started being employed, such as chaos theory and quantum theory, in an effort to create the "holy grail" of cryptographic systems and network security.

Quantum theory defines the non-classical qubit, which is the superposition of quantum states having no classical analog. It is also based on the "no cloning" and "no copying" theorem and on the Heisenberg's uncertainty. The quanto-mechanical properties of photons and the quantum mechanics provide a superb platform for communications security and particularly in secure optical communication networks, known as quantum cryptography and quantum networks, respectively.

Chaos is based on the particular behavior of certain non-linear functions, which for a minute change of parameters produce a huge and unstable output, known as the "chaotic regime". However, this chaos can be reproducible, which makes it attractive to secure communications. Thus, the combined Chaos and Quantum theories play a synergistic role in the quest for the "holy grail" in communications security.

In this talk we explain with simple examples quantum cryptography and protocols, as well as chaos and chaotic processes. We also describe the synergy of chaos and quantum theories as well as the increased efficiency and speed of the quantum key establishment.

BRIEF BIOGRAPHY

Stamatios V. Kartalopoulos, PhD, is currently the Williams Professor in Telecommunications Networking at the University of Oklahoma. His research emphasis is on optical communication networks (FSO, long haul, and FTTH), optical technology including optical metamaterials, and optical communications security including chaos, quantum cryptography and key distribution. Prior to this, he was with Bell Laboratories where he defined, led and managed research and development teams in the areas of DWDM networks, SONET/SDH and ATM, Cross-connects, Switching, Transmission and Access systems. He has received the President's Award and many awards of Excellence.

He holds nineteen patents in communications networks and has authored more than two hundred scientific papers, nine reference textbooks in advanced fiber optic communications and security, and has also contributed several chapters to other books.

He has been an IEEE and a Lucent Technologies Distinguished Lecturer and has lectured at international Universities, at NASA, conferences and research forums. He has been keynote speaker of major international conferences, has moderated executive forums, has been a panelist of interdisciplinary topics, and has organized symposia, workshops and sessions at major international communications conferences.

Dr Kartalopoulos is an IEEE Fellow, chair and founder of the IEEE ComSoc Communications & Information Security Technical Committee, board member of IEEE Potentials magazine, IEEE Press/Wiley Series editor, member at large of IEEE New Technologies Directions Committee, and has served editor-in-chief of IEEE Press, chair of ComSoc Emerging Technologies and of SPCE Technical Committees, Area-editor of IEEE Communications Magazine/Optical Communications, member of IEEE PSPB, and VP of IEEE Computational Intelligence Society.

INFORMATION SECURITY
The SCAN – Secure Processor with Crypto-Biometrics Capabilities

Nikolaos Bourbakis
Wright State University, U.S.A.

Abstract: Secure computing is gaining importance in recent times as computing capability is increasingly becoming distributed and information is everywhere. Prevention of piracy and digital rights management has become very important. Information security is mandatory rather than an additional feature. Numerous software techniques have been proposed to provide certain level of copyright and intellectual property protection. Techniques like obfuscation attempts to transform the code into a form that is harder to reverse engineer. Tamper-proofing causes a program to malfunction when it detects that it has been modified. Software watermarking embeds copyright notice in the software code to allow the owners of the software to assert their intellectual property rights. The software techniques discourage software theft, can trace piracy, prove ownership, but cannot prevent copying itself. Thus, software based security firewalls and encryption is not completely safe from determined hackers. This necessitates the need for information security at the hardware level, where secure processors assume importance.

In this talk the SCAN-Secure Processor is presented as a possible solution to these important issues mentioned above. The SCAN-SP is a modified Sparc V8 processor architecture with a new instruction set to handle image compression, encryption, information hiding and biometric authentication. A SCAN based methodology for encryption and decryption of 32 bit instructions and data and a Local-Global graph based methodology for biometric authentication is presented. The modules to support the new instructions are synthesized in reconfigurable logic and the results of FPGA synthesis are presented. The ultimate goal of the presented work is the tradeoffs that exists between speed of execution and security of the processor. Designing a faster processor is not the goal of the presented work, rather exploring the architecture to provide security is of prime importance.

BRIEF BIOGRAPHY

Nikolaos Bourbakis (IEEE Fellow) is an OBR Distinguished Professor of IT and the Director of the Assistive Technologies Research Center (ATRC) at Wright State University, OH. He pursues research in Applied AI, Machine Vision, Bioinformatics & Bioengineering, Assistive Technologies, Information Security, and Parallel- Distributed Processing funded by USA and European government and industry. He has published more than 330 articles in refereed International Journals, book-chapters and Conference Proceedings, and 10 books as an author, co-author or editor. He has graduated 17Ph.Ds and 37 Master students. He is the founder and the EIC of the International Journal on AI Tools, the Editor-in-Charge of a Research Series of Books in AI (WS Publisher), the Founder and General Chair of several International IEEE Computer Society Conferences, Symposia and Workshops, an Associate Editor in several IEEE and International Journals and a Guest Editor in 18 journals special issues. His research work has been internationally recognized and has won several prestigious awards. Some of them are: IBM Author recognition Award 1991, IEEE Computer Society Outstanding Contribution Award 1992, IEEE Outstanding Paper Award ATC 1994, IEEE Computer Society Technical Research Achievement Award 1998, IEEE ICTAI 10 years Research Contribution Award 1999, IEEE Symposium on BIBE Outstanding Leadership Award 2003, ASC Award for Assistive Technology 2005, University of Patras Degree of Recognition 2007.

e-Business Applications

FULL PAPERS

MOBILE MARKETING COMMUNICATIONS TO THE YOUTH
An Analysis of the MXit Platform

Charmaine du Plessis

Department of Communication Science, University of South Africa, Pretoria, South Africa
dplestc@unisa.ac.za

Abstract: This paper explores the various ways in which mobile marketing communications are available to the youth on a very popular South African mobile social network site known as MXit. The popularity of cellular phones has resulted in mobile-based brand advertising and promotions that specifically target the youth. The mobile medium is used by marketers as an additional channel to current marketing communications strategies. Because the mass market in South Africa does not use expensive cellular handsets, the format of mobile marketing communications is kept simple. MXit offers its subscribers effortless and accessible technology as well as appealing content, which is why it has become a prominent advertising platform. This is illustrated by examining two mobile marketing communication campaigns (one completed and one continuous) using MXit as an additional marketing communications channel to existing traditional marketing communications channels.

1   INTRODUCTION

The mobile platform has become increasingly dominant as a method of communication, simply because a cellular phone is permanently switched on, is within reach and is a very personal device. Marketers are therefore making use of the opportunity to communicate information about their products and services to their target market in what is known as "mobile marketing".

According to Pousttchi and Wiedemann (2006), mobile marketing is a division of mobile commerce. They define *mobile commerce* as "any kind of business transaction, on the condition that at least one side uses mobile communication techniques ". The focus of mobile marketing lies in the fourth P of the four Ps of the marketing mix: promotion, which is why they therefore prefer the term "mobile marketing communications". Pousttchi and Wiedemann (2006) define *mobile marketing* as a "form of marketing communications using mobile communication techniques to promote goods, services and ideas". This definition of mobile marketing is adopted for this paper.

A variety of technological platforms support mobile marketing applications, such as wireless application protocol (WAP), short message service (SMS) and multimedia message service (MMS)

(Carroll, Barnes, Scornavacca & Fletcher 2007:81).

Because a cellular phone is a very personal item, it has a great deal of potential in terms of tailor-made, personalised, two-way interaction, especially between the marketer and the customer. Other advantages include high reach, relatively low cost, high retention and flexibility. Marketers can virtually reach cellular phone users anytime and anywhere, even more so than the internet, because customers always carry their cellular phones with them. Marketing communication messages sent via a mobile medium represent mediated interactivity (communication between two individuals facilitated by a device). However, interactivity can only occur if the recipient of the message chooses to interact with the sender. Marketing communication messages should therefore be designed to generate interactivity as a result of its perceived value (Sinisalo, Salo, Karjaluoto & Leppaniemi 2007:774).

To develop more effective ways to target consumers, service providers have begun to explore the marketing potential of mobile social networks of consumer-to-consumer communication. For example, Vodafone pushes Multi-Media Messaging (an enriched form of texting that enables consumers to attach sounds, pictures, and videos to their text messages by enticing SMS users to share and experiment collectively with this new service

(Kleijnen, Lievens, De Ruyter & Wetzels 2009:15). Advances in internet tools and cellular phone features allow numerous mobile applications such as to personalised ring tones, phone interface and style, screen savers, information, SMS alerts, address books, pictures albums and MMS services. Because they can personalise their own mobile service experiences, cellular phone users increasingly demand affordable and reliable services that correspond exactly to their specific individual needs, lifestyles and preferences (Sigala 2006:395).

SMS is still the most popular mobile data application for marketing (Timpson & Troutman 2009:59), but marketers are also exploring other mobile marketing strategies to reach more consumers; in other words, not only those who have been granted permission to be opted in to receiving mobile advertising messages. Marketers increasingly let consumers subscribe to a branded mobile channel (contracted with a mobile service provider) where they can access different offerings that may include exclusive content as well as applications, games, special opportunities, incentives, and even emotional experiences to strengthen the value of the sponsoring brand while, at the same time, generating business value. As a result, organisations engage with consumers on a deeper level and increase their brand loyalty; consumer retention is also enhanced (Friedrich, Gröne, Hölbling & Peterson 2009: 54). Because a cellular phone usually belongs to one person only, mobile service providers are also able to do "contextual marketing" to their subscribers based on tracked information such as geographical position of users, technical addresses of the cellular phones and customer profiles. Contextual marketing refers to the extent to which e-businesses use the internet to provide customers with relevant information in the right context and in real-time (Lee, Korea & Jun 2007:799).

However, marketing via a mobile platform also poses various limitations and challenges, such as small screen size, limited bandwidth, device diversity, limited set of visual and audio capabilities, and issues such as permission-based regulations, consumer trust, and perception of personal intrusion. Marketers therefore consider the mobile medium as a complementary or substitute channel to current marketing communication strategies (Sinisalo, Salo, Karjaluoto & Leppaniemi 2007:775). Alternatively, they use mobile marketing communications primarily to strengthen the organisation's brand (Okazaki 2009:168).

2 MOBILE COMMUNICATIONS TO THE YOUTH

The popularity of cellular phones has resulted in mobile-based brand advertising and promotions that specifically target the youth. Although the youth market is often referred to by marketers as "adolescents", "Generation Next", "Generation Y" or even the "internet or iGeneration", researchers of the Golden Age youth survey – which was completed in 2008 – now challenge the traditional demographic definition of "youth". They argue that the traditional definition of youth is no longer relevant in today's society and that the youth market is much larger than is believed. Marketers should therefore "target consumers based upon their engagement and participation in youth culture rather than on their chronological age" (http://www.netimperative.com/news/2008/october/ 1st/study-debunks-2018golden-age-of-youth2019).

Cellular phones are essential to many youth lifestyles. Not only have they become indispensable fashion statements, but they also help their owners connect with and synchronise peer networks. Using cellular phones may also help young people gain and maintain peer group acceptance (Grant 2007: 224). Young people have an "intimate relationship" with their cellular phones. Mobile marketers realise that they can reach the youth at any time (Okazaki 2009). Numerous studies examining the reason for youth's rapid adoption of cellular phones refer to the uses and gratification theory, which assumes that media audiences are active, seek goal-directed gratification that the choice of media lies with the individual, and that media compete with other sources for consequent satisfaction (Leung & Wei 2000). Other studies refer to the social network theory, which stresses the importance of personal attributes in terms of obtaining a certain position in a network (Kleijnen et al 2009). Ling (2007:60) refers to research in Europe, which indicates that mobile communication enhances informal social interaction within the user's immediate group of friends – which makes cellular phones an indispensable item.

Research by Nielsen in 2009 on media usage by teenagers in the United States (and other countries in which the company tracks internet use, such as the United Kingdom, Italy, France, Spain, Germany, China, Brazil, Australia and South Africa) reveals that teenagers send text messages at incredible rates, and that they are also early adopters of all mobile media. Marketers can therefore reach them through other mobile platforms; in other words, not only

through SMS marketing (Nielsen, 2009). Another study by Nielsen in 2008 found that teen mobile media users were approximately three times as interested in mobile advertising as the total subscriber population. This means that more than half of teen mobile media users can be considered as open to mobile advertising (Nielsen, 2008).

Jenkins (2006) explains that, since young people are generally skilful when it comes to new technologies and therefore adopt new mobile technologies quite quickly, marketers also attempt to engage them with an organisation's brand via mobile databases. For instance, some magazines aimed at teenagers may encourage their readers to opt in to receiving free SMS messages (for instance, *CAR* and *Seventeen* magazines) to encourage interaction with the brand. When marketing to the youth, "safe and ethical communication and interaction" is the key to the success of any mobile communication. All communication should be brand related and integrated into the organisation's brand strategies (Jenkins, 2006).

Many young people in South Africa have adopted inexpensive mobile and rich-media and internet-based applications (Kreutzer, 2009). However, because the mass market in South Africa does not use expensive cellular handsets, the format of mobile marketing communications is kept simple to suit users' phones (http://www.totallymad.co.za/?idstory=20984).

3 RESEARCH QUESTION AND METHODOLOGY

The research question for this study is:
How do mobile marketing communications occur in a popular South African mobile social network? A case-study approach was used, because mobile marketing communications is still not fully understood. According to Rowley (2002:16), a case-study approach is appropriate when existing theory is lacking. A case-study approach is also useful when the researcher mainly asks "how" and "why" questions and uses a research strategy designed to investigate an event in its real-life context. A purposive sample of two case studies was selected (one completed and one continuous), because the two campaigns exemplify how mobile marketing communications take place within a popular mobile social network. These two case studies were then analysed using the following evaluation criteria (which were taken from the literature):

1. Which marketing communication tool (as identified in MXit) was used?
2. Which aspects of mobile marketing communications were adopted by using this tool?

In addition virtual ethnography was used which is the practice of ethnography but only in an online setting (Bosch 2009:187).

Various mobile social networks exist in South Africa (for instance, The Grid and Ubyou), of which the most popular, among the youth, is MXit.

4 MOBILE MARKETING COMMUNICATION USING THE MXIT PLATFORM

4.1 MXit Explained

According to Andrews (2008), MXit was introduced into South Africa in 2005 by a company known as MXit Lifestyle. It was introduced as a free mobile instant messenger programme tailored specifically for cellular phones – to provide users with instant access to social networks and music downloads. Chigona and Chigona (2008) explain that MXit is "network-independent" and can operate between users on any network provider; users do not have to be on the same network. Some South African youths also use other international MIMs (for instance, mig33 and Fling).

MXit has since released a PC Beta version to also include access to other features such as information, the gallery, Joe Banker, and Tradepost (see 4.3 below). MXit is available for virtually all Java phones and has recently launched the new MXit Elite for smart-phones, an MXit client for Blackberry, and is looking to launch an iphone application. This means that users can chat to other MXit users on their cellular phones or PCs anywhere in the world via the internet through GPRS or 3G, rather than through standard SMS technology. Users can chat using up to 1 000 characters at a time, at a fraction of the cost of a standard SMS. In addition to basic chat services, MXit offers users the chance to meet people, play games and customise their phones. The latest MXit version, V5.6.3, allows users to tab control for chat screens, make calls using MXit, and includes more text mark-ups and a faster start up and log in. MXit now has (2009) more than 14 million users worldwide (Andrews, 2008) (since its expansion into the international market).

MXit users send approximately 35 000 messages per second during peak times and also has a

community that visits the system more than 20 million times a day (http://www.bizcommunity.com/Article/410/78/379 52.html). This allows users to chat individually or in chat rooms, the only charge being the cost of the data sent, which makes it far cheaper to send messages via MXit than SMS) (Pasquinelli, 2009).

As from September 2009, MXit was also integrated with Facebook, allowing users to interact with their Facebook contacts via MXit's platform. MXit users can already access Yahoo! Windows Live Messenger, G Talk, AIM and ICQ (http://www.mobilemarketingmagazine.co.uk/2009/ 09/mxit-adds-facebook-integration.html).

Ramachandran (2009) explains that MXit's popularity can also be attributed to the fact that it accelerates social change and does not merely concentrate on music and games. For example, users can download books at a very low cost; there are education tools available for children (making it an ideal e-learning platform), and real-time drug counselling can also be done. It also provides inexpensive access to online services to people who do not have internet access. Additional services that make the service valuable to young people are a mathematics programme that provides online tutoring and quizzes (in partnership with the Department of Education, Nokia and a Finnish funding agency) and an application that helps young people to prepare for their driver's licence test (Ramachandran, 2009).

Other tutoring services on MXit are called *Imfundo Yami Imfundo Yethu* (which also provides assistance with mathematics) and Angel (health, AIDS, drugs) (Pasquinelli, 2009).

4.2 Popularity of the MXit Platform Among the Youth

MXit is extremely popular among South Africa's youth and was judged the "coolest" website and social networking service facility in 2008 and "the most frequently-used platform in the social media site" and website category in the *Sunday Times* 2009 Generation Next surveys, conducted by HDI Youth Marketers in partnership with the *Sunday Times*, and supported by Monash South Africa. The *Sunday Times* Generation Next Survey surveyed more than 5 000 urbanites in the 8 to 22 year old age categories (Andrews, 2008; HDI Youth Marketers, 2009).

A pilot study by Chigona, Chigona, Ngqokelela and Mpofu (2009), who conducted interviews with 12 young MXit users and nine parents, indicate that

the young people who were interviewed use MXit mainly for social networking. The MXit system also allows them to sustain and/or nurture networks of friends and associates. A quantitative study by Kreutzer (2009) of 500 grade 11 students in a low-income area of Cape Town confirms that 47% of the respondents use MXit on a typical day, making it the most frequently used social networking application in the sample. Reliable figures for mobile internet use in South Africa by the research company World Wide Worx indicate that there were 50 million mobile connections by the end of 2008, of which 68% represent individual users (Goldstuck, 2009).

Research was completed in South Africa in 2009 by the Bureau of Market Research's Youth Research Unit (YRU) (based at the University of South Africa) on new media usage among adolescents in selected schools in Tshwane (Pretoria). Findings reveal that tools such as the internet and cellular phones are constantly being used by youths to not only satisfy their own personal needs, but that they are willing to adjust to a new lifestyle where they use the latest media to their advantage. It also seems that 80% of cellular phone users in the sample use MXit for entertainment purposes, for help with their homework, to send interpersonal messages, and to make appointments. Of the 500 respondents in the study, 80% have access to the internet, while all have access to a cellular phone (Rapport, 2009).

4.3 Mobile Marketing Communications to the Youth using MXit

Because of its popularity among the youth, MXit has become a major advertising platform for reaching the South African youth market, some of whom have money of their own; it also exerts an influence on their parents in terms of how they spend their money. The primary target market for MXit is all males and females aged 16 to 25, while the secondary target market comprises all males and females aged 26 to 35. The 18 to 25 age bracket hold the lion's share (60%) of the MXit user market (http://www.marketingweb.co.za/marketingweb/vie w/marketingweb/en/page72308?oid=119872&sn=M arketingweb+detail).
The following types of mobile marketing communication tools are available on MXit:

4.3.1 Splash Screen Advertising

All MXit users have to register (for free), which provides advertisers with clear demographics in terms of the age, gender and location of their target

market. Splash screen advertisements are popular among advertisers. Full colour advertisements display for five seconds every time a user logs in. Users can only receive one splash advertisement at a time – similar to radio or television commercials. Splash screen advertisements are used to promote a particular portal to MXit's existing community, create brand awareness within the MXit community, and to complement an existing advertising campaign. Users can be targeted according to gender and age. Successful advertisers make use of striking graphics and send out a clear call to action (Fulford, 2008).

4.3.2 Hosted Brand Portals

According to Andrews (2008), part of the effectiveness and likeability of MXit's advertising is that it is all entirely opt-in or permission marketing, meaning users are not spammed with unnecessary, irrelevant advertising. If MXit users want information and services, they become part of the MXit world, since they then choose to add a brand as a personal contact on MXit. MXit users can become part of the advertiser's world, in which case they can interact directly with the marketer.

Users will add hosted brand portals listed under MXit Mix as a personal contact. The organisation then provides information and content via this contact (in XML format). Content is structured to complement existing media and also to increase traffic to existing media (Fulford, 2008). Various brands that are popular among the youth have hosted brand portals (for instance, Billabong and Samsung).

4.3.3 Hosted Chat Zone

Marketers can also create a virtual world where users can "live a brand". Branded and non-branded virtual items are made available for sale in the chat rooms. Users have the opportunity to chat with popular personalities. A competition can also be run to win an invitation to chat to a celebrity at a particular time. Treasure hunts can be created where users have to visit other media to receive clues for the chat rooms. A chat zone can have up to 100 chat rooms, with a maximum of seven users per room at any time. The advertiser has moderator status to initiate conversation and persuade users to chat about the brand. The chat rooms can be named to suit a particular brand, product, or campaign. The creator can also create his own virtual commands and enter any chat room (even if it is full), since he has "moderator" status for the rooms he creates (Fulford, 2008).

4.3.4 Targeted Two-way Communication

By using the database of users who subscribe to a particular service or chat room, organisations can engage in targeted, two-way communication by sending out questionnaires to a specific target market (Fulford, 2008).

4.3.5 Tradepost

Tradepost is a default contact on MXit. Users visit Tradepost to trade Moola (MXit's virtual currency) for chat rooms, skinz, contacts and premium services and content etc. Listing a contact on Tradepost makes it easy for users to add this contact. MXit provides one free Tradepost message to notify users that a new contact is available. Any person who adds a contact can be selected to win a sponsored prize. Additional messages can be purchased. The skinz are listed on Tradepost for sale to users, which enhances the advertiser's corporate identity and recognition. Advertisers can also sponsor Tradepost Treasure competitions, where MXit sends out a Tradepost message every day for up to seven days (where the advertiser is acknowledged as the sponsor of the competition). MXit's Skinz functions as personalised themes for MXit users. They can also be designed according to specific graphics, based on a brand's logo or a specific promotional campaign, and should be relevant and current (Andrews, 2008 & Fulford, 2008).

The trader keeps in regular contact with the user, providing information about the latest goodies and competitions available on MXit. Tradepost provides a list of available services, ranging from news reviews (both local and international) to movie reviews and times; horoscopes and more. Services are paid for in moola – one moola is equivalent to one cent, which can be purchased in partnership with Standard Bank by directly exchanging money for virtual currency. Tradepost also offers users emoticards, skinz, MXit games and role-playing games, Trivit (a general knowledge test), chat rooms and grown up chat rooms, geographical, topical, flirt, teen and celebrity chat zones as well as a dating game, weather updates and career tips (Andrews, 2008).

5 MOBILE MARKETING COMMUNICATION CAMPAIGNS USING MXIT AS AN ADDITIONAL CHANNEL

5.1 The Engen Endless Summer Campaign

The most successful mobile marketing communications campaign up to date using MXit is no doubt the energy company Engen's Endless summer campaign, which was launched in December 2006 for the duration of the summer holidays. Engen is an Africa-based energy company with a focus on the downstream refined petroleum products market and related businesses. Engen's core business is the refining of crude oil and marketing of their primary refined petroleum products, and the provision of convenience services through an extensive retail network (Engen Corporate Report, 2009).

5.1.1 Types of MXit Marketing Communications Tools used

The aim of the Endless Summer campaign was to entertain young people sitting on the back seat of their parents' cars en route to holiday destinations and to get parents to stop at the next Engen convenience store. The target market was teenagers aged 11 to 18. Splash screen advertising was used as well as a hosted chat zone. Keywords were placed on billboards throughout the summer holidays along all the main holiday routes. Each keyword entitled users to a different download from mobile wallpapers to ringtones and games by using virtual money. These keywords needed to be entered into the Endless Summer chat zone or mobile site. The only mention of what to do with the keywords was on MXit, the mobile chat service and not on the billboards. This spread the viral element: young people knew what the keywords meant, but their parents did not. Chat room moderators spoke to young people in the chat rooms, motivating them to do more downloads. Secret keywords were also provided at certain times.

The success of the Engen campaign reiterates the findings of the Nielsen 2008 study, namely, that young mobile users can be considered open to mobile advertising (Nielsen, 2008). Given that all marketing communications were brand related and integrated into Engen's brand strategy (see Jenkins 2006), the target market could personalise the Engen

brand experience.

See table 1 below for the types of MXit marketing communication tools used:

Table 1: Types of MXit marketing communication tools used for the Engen Endless Summer campaign.

Type of MXit marketing communication tool	Aspects of mobile marketing communications adopted
Splash screen advertising	Integration with Engen's brand strategy to get young people to visit other portals in MXit; targeted tailor-made communication.
Tradepost	Personalisation of the Engen brand; strengthened the value of the sponsoring brand and, at the same time, generated business value by the revenue earned from the sales of rich media content.
Hosted chat zone	Engagement with the Engen brand; increased brand loyalty; opportunity for subscribers to "live" the brand.
Hosted brand portal	Engaged with consumers on a deeper level, increased brand loyalty and enhanced consumer retention.

5.1.2 Results of the Engen Endless Summer Campaign

Ten thousand young people added the *EndlessSummer* contact as a branded channel contact in the MXit platform. Over three million messages were posted in the chat zone. More than 12 000 downloads were done in six weeks. This campaign also won gold in the annual Loerie awards for best advertising in the category digital mixed media campaign in 2007.

5.2 *CAR* Magazine

Published by Ramsay media automotive, CAR magazine is Southern Africa's foremost multi-media automotive consumer brand with a readership of 1 022 000 for the printed magazine by the end of 2009. In addition to this, by the end of 2009, the cartoday.com website had a readership of 59 585 unique visitors, mobile.cartoday.com had a readership of 5 850, the CAR branded channel on MXit already had 99 203 subscribers, its Facebook profile had 2 259 friends and it had 559 followers on Twitter. Rich content can also be watched on YouTube and myvideo

(http://www.ramsaymedia.co.za/sections/brands/car/index.asp).

Sixty percent of its readers are young males under the age of 35. The magazine has also been nominated as the winner in the Coolest Magazine and Coolest Male Magazine categories 2009 in the *Sunday Times* Generation Next Survey (http://www.mediaupdate.co.za/?IDStory=16423).

5.2.1 Types of MXit Marketing Communications Tools used

A key objective of the marketing strategy of CAR magazine is to ensure that its content stays relevant to its young readers, who are also technologically well-informed. Apart from the magazine's mobi-site, which has an average of 70 000 page views and 10 000 unique users per month, it has a hosted brand portal on MXit, where users can add them as a contact by paying a minute fee (http://www.ramsaymedia.co.za/sections/brands/car/index.asp).

Within the hosted brand portal, CAR magazine uses its own rich media content to supplement content in other media, or runs a mobile campaign for a client that will simultaneously benefit the *CAR* brand. For instance, in June 2009, it developed content for the Golf 6 launch campaign. During this campaign *CAR* magazine distributed branded Golf 6 videos, ringtones and wallpapers to thousands of subscribers across *CAR*'s mobile platforms, including the *CAR* mobi-site and the magazine's branded channels on MXit, MTN Loaded, as well as the Samsung Fanclub and Nokia WAP portals (http://www.mediaupdate.co.za/?IDStory=18632). The magazine also ran an integrated campaign for a top South African band known as the Parlotones. In October 2009 *CAR* magazine teamed up with the Parlotones to record a viral video of the band's hit *Push Me to the Floor*, which was made available on MXit and on the *CAR* website. Mobile content subscribers could then enter a competition in the October issue of the *CAR* magazine where they had to identify the mystery driver in the video to win amazing Parlotones prizes (http://www.marketingmix.co.za/pebble.asp?relid=8403).

See table 2 below for the types of MXit marketing communications tools used:

Table 2: Types of MXit marketing communications tools used for *CAR* magazine.

Type of MXit marketing communications tool	Aspects of mobile marketing communications adopted
Tradepost	Personalisation of the *CAR* brand; strengthens the value of its brand and, at the same time, generates business value by revenue earned from sales of rich media content.
Hosted brand portal	Strengthens the *CAR* brand by integrating traditional media with that of mobile media. Constantly engages with consumers on a deeper level, increases brand loyalty and enhances consumer retention. Runs mobile campaigns for clients that create better brand awareness.

CAR magazine continuously uses mobile marketing communications as a supplementary channel primarily to strengthen its brand (Okazaki 2009: 168). By getting its consumers to subscribe to a branded mobile channel in MXit – where they can access various offerings such as exclusive content as well as applications, games, special opportunities, incentives, and having emotional experiences through brand interaction – the value of the *CAR* brand is enhanced and business value is also generated (Friedrich et al 2009: 54).

5.2.2 Some Results of the *CAR* Magazine MXit Campaigns

By the end of 2009, subscribers to CAR's hosted brand portal on MXit was fast approaching the 100 000 mark (a 400% growth since its launch in October 2008). In addition, more than 250 000 units of CAR's rich media content (ranging from wallpapers to ring tones) have been sold since the launch. This clearly indicates that MXit can be a profitable marketing communications tool for advertisers who produce MXit friendly content.

In June 2009, during the launch of the Golf 6 campaign, a total number of 98 765 subscribers to the *CAR* hosted brand portal on MXit had access to the Golf 6 videos, exhaust tones and wallpapers. Other media exposure included 15 560 unique users on mobile.cartoday.com, while content was also available on *CAR's* channels – on MTN Loaded, Samsung Fanclub and Nokia WAP portals. There were some 2890 views of the branded Golf 6 video

on *CAR*'s YouTube channel, which includes a link to the CARtoday site, and 1 289 views of the clip on myvideo (http://www.marketingmix.co.za/pebble.asp?relid=8403).

6 CONCLUSIONS

A successful mobile social networking site among the youth, known as MXit, has become a powerful advertising medium in South Africa owing to its simple, accessible and affordable technology. By providing inexpensive, accessible and relevant content, MXit has succeeded in increasing the number of its subscribers, by the end of 2009, to more than 14 million.

The findings indicate that more South African youth are using mobile phones to connect with their friends or peers and to access digital content repeatedly. Although mobile marketing communication that specifically targets the youth in South Africa is still in its formative years, marketers are gradually beginning to tap into this lucrative market. By keeping access to mobile content relatively inexpensive, both audience reach and repeat purchases are encouraged. A mobile campaign needs to be both topical and interesting t young MXit users in order to be successful.

The two case studies explained in this paper differ in terms of MXit tools used. In the completed Engen case study the marketers focused on all the MXit tools available because of the high percentage of young people who are using MXit in South Africa. They could reach their young target audience instantly by keeping them interested in the tailor-made marketing communication messages. In the case of the continuous CAR magazine mobile campaign, the marketing strategy is rather to direct the reader to the print publication. In this case MXit is only used as a complementary tool to traditional media and only the most suitable tools for this purpose are being used for brand building purposes.

While the aspects described in this paper have been practiced in Europe and other countries for some period of time now, mobile marketing in South Africa, a third world country, has only augmented in recent years. Future studies could therefore include how South Africa currently compares with first world countries in terms of mobile marketing or how mobile marketing using MXit as a platform differs from marketing using other social media platforms. South African marketers use various MXit marketing communications tools primarily to strengthen their brands and to earn revenue through selling rich-media content to loyal users, which personalises the brand experience.

REFERENCES

Andrews, L., 2008. MXit as a mobile media advertising platform.Homepage: http://www.mediaupdate.co.za/?IDStory=13405#.

Bosch, T., E., 2009. Using online social networking for teaching and learning: Facebook use at the University of Cape Town. *Communicatio* 35(2): 185-200.

Carroll, A., Barnes, SJ., Scornavacca, E., Fletcher, K., 2007. Consumer perceptions and attitudes towards SMS advertising: recent evidence from New Zealand. *International Journal of Advertising*. 26(1): 79 -98.

Chigona, W., Chigona., A, Ngqokelela., B & Mpofu S., 2009. MXIT: Uses, Perceptions and Self-justifications. *Journal of Information, Information Technology, and Organizations*. Homepage: http://jiito.org/articles/JIITOv4p001-016Chigona369.pdf.

Chigona, A., & Chigano W., 2008. MXit it up in the media. Media discourse analysis on a mobile instant messaging system. *The Southern African Journal of information and Communication*. 9: 42- 57.

Engen Limited corporate report, 2009. Homepage: http://www.engen.co.za/downloads/media_centre/publications/Engen%20Corporate%20Report%202009.pdf.

Friedrich, R., Gröne, B, Hölbling., K, Peterson., M. 2009. The March of Mobile Marketing: New Chances for Consumer Companies, New Opportunities for Mobile Operators. *Journal of Advertising Research*. March. 54 – 61.

Fulford, S., 2008. MXit Business Presentation. Homepage: http://www.alphabetsoup.co.za/intestestingcontent/Business%20PresentationJune%202008%20%5BCompatibility%20Mode%5D.pdf.

Goldstuck, A., 2009. Mobility survey. Homepage: http://www.worldwideworx.com/archives/204.

Grant, I., 2007. Why young consumers are not open to mobile marketing communication. *International Journal of Advertising*, 26(2): 223–246.

HDI Youth Marketeers. 2009. Generation Next: A new style of cool. 1 June. Homepage: http://www.bizcommunity.com/Article/196/347/36512html.

Jenkins, F., 2006. Mobile marketing. *Young Consumers*. Quarter 1: 60-63.

Kleijnen, M., Lievens, A., de Ruyter, K & Wetzels, M., 2009. Mobile Services: Knowledge Creation Through Mobile Social Networks and Its Impact on Intentions to Use Innovative. *Journal of Service Research*. August 12(1): 15-35.

Kreutzer, T., 2009. Internet and online media usage on mobile phones in low-income urban youth in Cape Town, Paper submitted to the pre-conference workshop of the 2009.

Lee, TM., Korea, C & Jun, JK., 2007. Contextual perceived value? Investigating the role of contextual

marketing for customer relationship management in a mobile commerce context. *Business Process Management Journal*. 13(6): 798-814.

Leung, L., & Wei, R., 2000. More than just talk on the move: uses and gratifications of the cellular phone. *Journalism and Mass Communication Quarterly*, 77(2):308–320.

Ling, R., 2007. Children, youth, and mobile communication. *Journal of Children and Media*. 1(1).

Nielsen, 2009. How Teens Use Media. A Nielsen report on the myths and realities of teen media trends. June Homepage: http://blog.nielsen.com/nielsenwire/reports/ nielsen_howteensusemedia_june09.pdf. , Accessed on 4 September 2009.

Nielsen. 2008. Critical mass. The worldwide state of the mobile web. Homepage: http://www.nielsenmobile. com/documents/CriticalMass.pdf.

Okazaki, S., 2009. The Tactical Use of Mobile Marketing: How Adolescents' Social Networking Can Best Shape Brand Extensions. *Journal of Advertising Research*. March: 12 – 26.

Pasquinelli, E., 2009. MXIT as a case study: educational purposes colonize non-educational practices, and technologies. 18 September**.** Homepage: http://www.interdisciplines.org/mobilea2k/papers/3/3.

Pousttchi, K., Wiedemann, DG., 2006. A Contribution to Theory Building for Mobile Marketing: Categorizing Mobile Marketing Campaigns through Case

Study Research. 20 November. Homepage: http://mpra.ub.uni-muenchen.de/2925/1/MPRA_paper _2925.pdf.

Ramachandran, S., 2009. MXit Mixes Mobile Networks with Social Conscience. 18 September 2009. Homepage:http://www.businessweek.com/technology/ content/aug2009/tc20090820_669558.htm.

Rowley, J., 2002. Using case studies in research. *Management research news,* 25(1): 16 -27.

Sinisalo, J., Salo, J., Karjaluoto, H & Leppa¨niemi, M., 2007. Mobile customer relationship management: underlying issues and challenges. *Business Process Management Journal*. 13(6):771-787.

Sigala, M., 2006. Mass customisation implementation models and customer value in mobile phones services. Preliminary findings from Greece. *Managing Service Quality*. 16(4):395-420.

Technology is a powerful tool in marketing aimed at the youth. Rapport newspaper. 31 May 2009.

Timpson, S., & Troutman, M., 2009. The importance of a layered privacy policy on all mobile internet sites and mobile marketing campaigns. *International Journal of Mobile Marketing*. June. 4(1): 57 -61.

PREFERENCES OF HANDWRITING RECOGNITION ON MOBILE INFORMATION SYSTEMS IN MEDICINE
Improving Handwriting Algorithm on the Basis of Real-life Usability Research

Andreas Holzinger, Martin Schlögl

Institute of Medical Informatics, Statistics and Documentation, Research Unit HCI4MED, Medical University Graz Graz, Austria
andreas.holzinger@medunigraz.at

Bernhard Peischl

Institute of Software Technology, Graz University of Technology, Graz, Austria
bernhard.peischl@ist.tugraz.at

Matjaz Debevc

Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Austria
matjaz.debevc@uni-mb.si

Abstract: Streamlining data acquisition in mobile health care in order to increase accuracy and efficiency can only benefit the patient. The company FERK-Systems has been providing health care information systems for various German medical services for many years. The design and development of a compatible front-end system for handwriting recognition, particularly for use in ambulances was clearly needed. While handwriting recognition has been a classical topic of computer science for many years, many problems still need to be solved. In this paper, we report on the study and resulting improvements achieved by the adaptation of an existing handwriting algorithm, based on experiences made during medical rescue missions. By improving accuracy and error correction the performance of an available handwriting recognition algorithm was increased. However, the end user studies showed that the virtual keyboard is still the overall preferred method compared to handwriting, especially among participants with a computer usage of more than 30 hours a week. This is possibly due to the wide availability of the QUERTY/QUERTZ keyboard.

1 INTRODUCTION AND MOTIVATION FOR RESEARCH

In cases of emergency, rapid patient information collection is very important. This information is most often collected by first aiders (first responders) and paramedics (e.g. Red Cross). Prompt and accurately recorded and well communicated vital patient data can make the difference between life and death (Holzman, 1999), (Anantharaman & Han, 2001).

The data acquisition should have as little disruptive effect on the workflow of the emergency responders (rescue staff) as possible. A possible solution for data input can be an mobile application on a lightweight handheld device (Baumgart, 2005), (Chittaro, Zuliani & Carchetti, 2007).

Due to the fact that emergencies are usually within difficult physical situations, special attention to the design of information technology for emergencies has to be taken into consideration (Klann et al., 2008). A key issue of any such information system is the acquisition of textual information. However, extensive text entry on mobile devices is principally to be avoided and a simple and easy to use interface, in accordance with the proverb: less is more, is a supreme necessity (Holzinger & Errath, 2007).

The basic evidence that entering data onto a mobile device via a stylus is slower, more erroneous and less satisfactory for end users than entering data via a QWERTZ (de) or QUERTY (us) keyboard has been demonstrated in some studies (Haller et al., 2009), although, on the other hand the use of a stylus is much faster and more accurate than using finger touch (Holzinger et al., 2008b). A specific study for "Ambulance Run Reporting" shows good results for acquiring text with a virtual keyboard, while acquiring text by the application of handwriting recognition showed some serious usability problems (Chittaro et al., 2007). Motivated by this previous work, we focus in this work on handwriting recognition and on how to improve its usability – in case of need, also by adaptation of existing handwriting algorithms. Consequently, in this paper we report on real-life experiences and on some improvements achieved by the adaptation of an existing handwriting engine.

2 BACKGROUND

A big difficulty of handwriting recognition is that handwritten characters are variable on an individual basis and that these characters are usually separated into alphabets, numerals, and symbols, despite the different characters of the language itself. Although handwriting recognition will benefit in future from improved adaptive and context-sensitive algorithms, improving the user experience of novice end users with the respective technology is possibly the most important factor in enhancing user acceptance (MacKenzie & Chang, 1999). This is even more important in medical or health care contexts, where the difficulty is in the environmental conditions, e.g. if the person is on the move or in a hurry (Holzinger et al., 2008a). Whereas the first problem might be solved by the training modus opportunities, in order to adapt the system to the individual handwriting style, the second problem is only solvable by an extremely robust and usable system. Especially in the health care domain, good end user acceptance and usability can only be obtained by providing simple operation (good user guidance), very short response times and low error rates (Holzinger, Geierhofer & Searle, 2006).

Basically, there are several methods for handwriting recognition; these belong basically to two distinct families of classification:

I) Structured and Rule Based Methods

Because of the fuzzy nature of human handwriting, it makes sense to adapt the well known fuzzy logic technique for this purpose (Gader et al., 1997). Rather than evaluating the two values as in digital logic, fuzzy terms admit to degrees of membership in multiple sets so that fuzzy rules may have a continuous, rather than stepwise, range of truth of possibility. Therefore non-identical handwritten numerals, from same or different users, can be approximated using fuzzy logic for fast and robust handwriting recognition (Shi & Li, 2006).

II) Statistical Methods

a) Hidden Markov Modeling (HMM)

The attractiveness of HMM for various pattern recognition tasks is mainly due to their clear and reliable statistical framework. Many efficient algorithms for parameter estimation and model evaluation exist, which is an important prerequisite for their practical implementation for real-life applications (Plotz & Fink, 2009). The methods using HMM (Marti & Bunke, 2002), are based on the arcs of skeleton graphs of the words to be recognized and an algorithm applied to the skeleton graph of a word extracts the edges in a particular order, which is transformed into a 10-dimensional feature vector. Each of these features represent information about the location of an edge relative to four reference lines, the curvature and the degree of the nodes incident to the considered edge. Training of the HMM is done by use of the Baum-Welch algorithm, while the Viterbi algorithm is used for recognition (Bunke, Roth & Schukattalamazzini, 1995), (Xue & Govindaraju, 2006).

b) Neural Networks

The methods based on Neural Networks were driven by the emergence of portable, pen based computers. A typical approach is to combine an artificial neural network (ANN), as a character classifier, with a context-driven search over segmentation and word recognition hypotheses (Yaeger, Webb & Lyon, 1998).

However, handwriting recognition not only consists of the recognition itself; the data must undergo some preprocessing:

(I) Reduce noise;
(II) Normalization, and
(III) Segmentation.

The last step, the segmentation phase, segments the input into single characters (Plamondon & Srihari, 2000). Writing discrete characters requires no segmentation; this is done by the users themselves (Tappert, Suen & Wakahara, 1990).

Another way to improve recognition is to decrease the set of possible alternatives, such as to restrict the set to accepting only lower case letters or digits (Frankish, Hull & Morgan, 1995).

3 RELATED WORK

To date only a few studies considered handwriting recognition on mobile devices and very few in the health care domain.

A very early work by Citrin et al. report very general on the usage of a pen on a flat surface of a LCD unit (scribing and tapping). They reported that with the maximum rate of 100 selections of direction per second for pen, scribing may produce strokes with the speed of 300 (100×3) bps. However, no more results were found (Citrin et al., 1993).

MacKenzie showed that the recognition accuracy for a set containing upper and lower case letters was lower than for a set containing just lower case letters (MacKenzie et al., 1994).

Chittaro evaluated a system for recording data on a system during a running ambulance drive, having first responders as participants. Text entry via virtual keyboard and handwriting recognition (MS Transcriber – Calligrapher) were also performed. Text entering by handwriting was considered very laborious and difficult by the users (Mean 3.8, Var 6.6), while entering text by use of the virtual keyboard was quite easy (Mean 7.2, Var 1.8). (0=Hard, 9=Easy). Furthermore, they emphasized the bad usability of entering text by using handwriting recognition. Most words were wrongly recognized and there were enormous problems in correcting those wrongly recognized words (Chittaro et al., 2007).

4 METHODS AND MATERIALS

The aim of our study was to increase the performance of available handwriting recognition by improving accuracy and error correction following solid usability engineering methods (Holzinger, 2005).

We focused on separate character recognition, since the correction of a single letter, at the moment of false recognition, can be made more naturally, and efficiently, than attempting to correct or delete a single letter within a recognised word.

Due to limited space, there could be some problems inputting long words. Therefore, only one character at a time can be written and recognized.

4.1 Experimental Device

The device used for the prototype was an Asus MyPad A626 PDA (Personal Digital Assistant).

This device is equipped with an anti-glare touch screen display. For typing on the touch screen, a stylus is used.

Table 1 contains the technical specifications of this device.

Table 1: Specifications of the PDA ASUS MyPal A626.

CPU	Marvell XScale, 312MHz
Operating System	MS Windows® Mobile™ 6
Memory	256MB Flash ROM and 64 MB SDRAM
Display	3.5" Brilliant TFT LCD 65k full-colours, anti-glare 16-bit display QVGA, 240x320 px touch screen
Weight	158g
Physical dimensions	117 mm x 70.8 mm x 15.7cm

4.2 Dialog Design
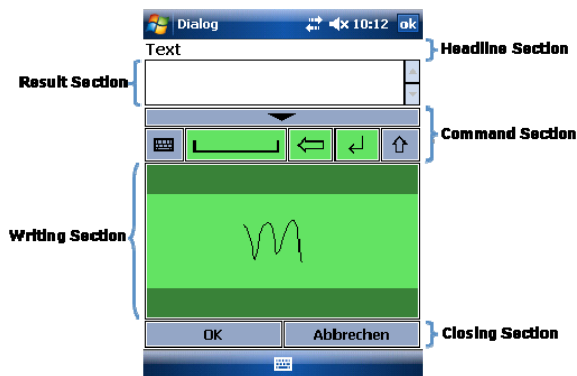
Figure 2: Design of the handwriting dialog.

The light green area within the writing sections defines the optimal size for handwritten lowercase characters of 80 points (Phatware, 2002).

4.3 Handwriting Recognition

We used the SDK of the handwriting recognition engine Calligrapher (in MS Windows® Mobile Transcriber) in the version 6.0 (Phatware, 2002).

This SDK makes it possible to define single character recognition. We can handle the results and a custom timeout (after which time the recognition starts) can be defined.

4.3.1 Adaptive Timeout

A handwritten character consists of one or more strokes. The recognition starts after the character is finished. The system has to await a timeout before starting recognition because the system doesn't know whether the character consists of just one or more strokes.

A stroke is defined as the writing from pen down to pen up (Tappert et al., 1990).

Because of the different writing speeds of each user, this timeout has to be calculated for each user. Therefore, the system stores the last ten times which elapse between two strokes.

$$T = \frac{\sum_{i=1}^{11} s(i)}{11} * \frac{X}{100}$$

Figure 3: Calculation timeout T [sec].

Figure 2 shows how the timeout is calculated every time a timeout is requested. s(1) is the last calculated average time between strokes, s(2)..s(11) are the last ten stored times between strokes. X is a factor, in this experimental setting X is 200. The result T is the timeout in seconds.

4.3.2 Correction Intervention

Calligrapher SDK 6.0 doesn't adapt recognition on users' handwriting because of the use of static Fuzzy-Neuronal Nets (Strenge, 2005).

There are problems with some user's style of writing letters – the user writes a letter (e.g. an "a") but the recognition engine recognizes another letter (e.g. figure 4).

Figure 4: Written "a" but not recognized as "a", instead as "ir".

A recognition result is a list of possible characters and its weight (maximum 5 entries). Every time the same letter is wrongly recognized for a user (as in Figure 2), the lists returned by the recognition are similar.

These lists (characters and its weight) with its representing letter are stored. Each of them is called schema.

During writing, the recognition result will be compared to the stored schemas as follows. (Example in Figure 5)
For each stored schema:
Characters from the result list and the list of the schema are compared. If the result list consists of 2 or 3 characters, at least 2 have to match to the stored schemas lists characters. (2 of 2, 2 of 3). If there are 4 or 5 characters in the result list, at least 3 have to match (3 of 4, 3 of 5). This means, the resulting list is validated to the list of the schema.

If the list is valid according to the list of the schema, the average deviation between these matching characters is calculated.

Inputted list		Stored Schema's list		Deviation
u	52	k	41	4
n	47	M	35	0
k	37	n	31	16
M	35	h	26	
A	22	m	22	
Validity: 3/5 → VALID				
Average Deviation (16+0+4)/3:				6

Figure 5: Example of a list comparison.

The representing letter of the schema with the lowest average deviation will be put in first place of the recognition result.

4.3.3 Calibration

The calibration is designed to collect user specific data for each letter. This data contains weights, which present every character explicitly. Also, schemas of wrongly recognized letters (Chapter 3.3.2) are collected.
The system prompts the user to input a letter.
If the result list of the recognition has the prompted letter in first place, the weight will be stored for this letter. In the calibration phase, at least 2 weights will be stored for each letter.

If not, the result list will be stored as a schema with the prompted letter as a representing letter. In the calibration phase, a maximum of 10 schemas for each letter is stored.
This calibration is done once for each user. A continuous calibration is also done during writing in the handwriting recognition dialog, saving weights and schemas for correctly recognized letters (but not for deleted letters)

4.3.4 Other Interventions on Recognition Results

To avoid side effects, the intervention described in Chapter 3.3.2 is only made when the weight of a recognized letter is less than the average weight for this letter (average of the weights for this letter collected by calibration).

Other interventions are made to avoid potential. problems with highly confusable pairs such as "r" and "v" (Frankish et al., 1995). (I) While writing a word, only letters and punctuation marks are valid, recognized results. (II) Just deleted letters (with BACKSPACE) are not valid, recognized results for the next recognition (III) Special handling for "O" and "0" as first letter of a word or number

4.4 Experiment

The real life environment is mostly a seat in an ambulance car. To avoid negative effects on ambulance responder's work, the experiment is done in their recess in the ambulance service rooms, simulating the circumstances (sitting in a car) by doing the experiment sitting on a chair, holding the PDA in their hand, without laying down the elbows on e.g. an armrest. (Kjeldskov et al., 2004) shows that simulating environments gives almost the same results.

Figure 6: Participants during experiments in real life.

Participants were people who work as ambulance officers (professionals, volunteers and former civilian service). No previous experience with mobile computers was required.

They were asked to fill out a background questionnaire to obtain data about their age, education and use of computers.

The prototype for the experiment is divided into two parts, one for virtual keyboard based text input, and the other for handwriting recognition input. Within these two parts, the users have the opportunity to

become familiar with the input methods. After that, the user has to input a given text to the experimental dialog (for measuring the accuracy). Due to measuring the accuracy, text entry is done as text copy (MacKenzie & Soukoreff, 2002). This text consists of 13 German words (94 characters without spaces, 106 with spaces). After the keyboard based experimental dialog, the calibration of the handwriting is done.

Speed in wpm, words per minute (Lewis, 1999, MacKenzie et al., 1994) and the accuracy of the handwriting recognition are measured and calculated. At the end, a feedback questionnaire is filled out by the user. Some questions are based on the study of Chittaro (Chittaro et al., 2007).

5 RESULTS

5.1 Participants

The participants of the experiment were professional (9) and volunteer (8) first responders of the Austrian Red Cross, one student of medicine and three others (because everyone could be a volunteer first responder).

10 are experienced on a PDA or a mobile phone with touch screen, while 11 have no experience with touch screens.

Their ages ranged from 20 to 85 years. Two elderly people (68 and 85 years) were chosen because they had never before used a QWERT keyboard or a PC.

The average use of a PC is 12.3 years, using a PC 31 hours per week. 11 participants use a PC ≤ 30 hours a week, while 10 participants use a PC for more than 30 hours.

One of the 21 participants was left-handed.

5.2 Accuracy

Overall		≤ 30 weekly usage		> 30 weekly usage	
Mean	Var	Mean	Var	Mean	Var
99.1	6.28	100	11.5	99.06	1.44

Figure 6: Accuracy inputting text with virtual keyboard [%]; all participants, participants ≤ 30 hours and above.

Overall		≤ 30 weekly usage		> 30 weekly usage	
Mean	Var	Mean	Var	Mean	Var
89.25	34.3	91.43	30.20	88.00	37.34

Figure 7: Recognition accuracy [%] of handwriting recognition; all participants, participants ≤ 30 hours and above with interventions.

Overall		≤ 30 weekly usage		> 30 weekly usage	
Mean	Var	Mean	Var	Mean	Var
84.66	57.6	86.99	79.15	83.33	38.21

Figure 8: Recognition accuracy [%] of handwriting recognition; all participants, participants ≤ 30 hours and above without interventions.

The participants using a PC ≤ 30 hours a week include the two elderly people.

The 85 year old participant has an accuracy of 89.2% for inputting text with the virtual keyboard and a recognition accuracy of 80.1% with interventions and 65.6% without interventions.

The 68 year old participant had an accuracy of 100% for inputting text with the virtual keyboard and a recognition accuracy of 95% with interventions and 90.8% without interventions.

The 85 years old participant has an accuracy of 89.2% for inputting text with the virtual keyboard and a recognition accuracy of 80.1% with interventions and 65.6% without interventions.

The 68 year old participant has an accuracy of 100% for inputting text with the virtual keyboard and a recognition accuracy of 95% with interventions and 90.8% without interventions.

5.3 Speed

Overall		≤ 30 weekly usage		> 30 weekly usage	
Mean	Var	Mean	Var	Mean	Var
13.17	27.7	12.88	29.46	13.43	18.29

Figure 9: Words per minute virtual keyboard; all participants, participants ≤ 30 hours and above.

Overall		≤ 30 weekly usage		> 30 weekly usage	
Mean	Var	Mean	Var	Mean	Var
8.44	4.59	8.11	5.37	8.71	1.95

Figure 10: Words per minute handwriting recognition; all participants, participants ≤ 30 hours and above.

Participants using a PC ≤ 30 hours a week include two elderly people.

The 85 year old participant wrote 2.87 wpm with the keyboard and 2.82 wpm with handwriting

recognition. The 68 year old participant wrote 4.88 wpm with the keyboard and 4.17 wpm with handwriting recognition.

5.4 User Questionnaire

Overall	Mean	Var
Keyboard		
Inputting Data (+4=easy, -4=difficult)	3.0	2.6
Correction of wrong inputted data (+4=easy, -4=difficult)	4.0	2.8
Handwriting		
Inputting Data (+4=easy, -4=difficult)	2.0	4.9
Correction of wrongly input/recognized data (+4=easy, -4=difficult)	3.0	1.4
Did the recognition slow down your writing (+4=no, -4=yes)	0.5	9.2
I would prefer (+4=handwriting, -4=keyboard)	-2.5	7.9
Basic Information		
Use of colour is (+4=useful, -4=useless)	2.0	2.5
The handwriting recognition positively surprised me (+4=yes, -4=no)	2.5	7.1
Characters on the PDA are easy to read (+4=yes, 4=no)	4.0	3.9

Figure 11: Overall results of user questionnaire.

Weekly computer usage [hours]	Mean (<=30)	Var (<=30)	Mean (>30)	Var (>30)
Keyboard				
Inputting Data	3.5	1.4	3.0	3.7
Correction of wrongly input data	4.0	1.7	3.5	4.5
Handwriting				
Inputting Data	2.5	4.2	0.5	3.8
Correction of wrongly input/recognized data	4.0	0.5	2.5	1.8
Did the recognition slow down your writing	2.5	8.4	-0.5	8.2
I would prefer	-1.5	2.0	-2.5	4.9
Basic Information				
Use of colour	2.0	2.5	1.5	2.6
The handwriting recognition positively surprised me	4.0	7.7	0.0	4.7

Characters on the PDA are easy to read	4.0	4.6	4.0	3.2

Figure 12: Results of user questionnaire for weekly usage of computer ≤ 30 hours and above.

6 CONCLUSIONS

Entering text with the virtual keyboard (Mean 3.0, Var 2.6) was easier for the participants than with handwriting (Mean 2.0, Var 4.9). However, compared to the study of (Chittaro et al., 2007), we could reach an significant improvement by inputting data with handwriting. Interestingly, inputting data by handwriting recognition was rated easier by participants who use computers less than or equal to 30 hours a week than by participants with extensively more use (Mean 2.5; Var 4.2; against Mean 0.5, Var 3.8 of virtual keyboard). Also, the correction on the handwriting recognition dialog was rated easier (Mean 4.0, Var 0.5; against Mean 2.5, Var 1.8; of virtual keyboard). Participants with a computer usage of more than 30 hours a week preferred the virtual keyboard (Mean -2.5, Var 4.9) more than the other participants (Mean -1.5, Var 2.0). This could be a result of hardly any handwriting during work and much more typing text on classical keyboards (QWERTZ or QUERTY). Consequently, the two elderly participants were included in this study, in order to obtain data regarding participants who never used any computer or handheld device. The elderly participants were the only ones who provided a complete preference to the handwriting recognition in contrast to the virtual keyboard. This is also clearly visible in the results for these participants, although both groups have quite comparable results in wpm for the virtual keyboard and the handwriting text input.

This is an interesting result; however, it is not of practical relevance, since there are hardly any people left – at least amongst people able to volunteer as a first responder – without experience on computer keyboards. Today, from elementary school on, children get used to work with computers by using the QWERTZ or QUERTY keyboard.

Nevertheless, our interventions on the basis of the results of the handwriting recognition, finally paid off in an significant improvement on the recognition accuracy (over all participants a better accuracy of Mean +4.39%, Var 9.54).

These interventions can also be useful for the improvement of other handwriting recognition engines, due to the fact that our interventions were only made on the results of the engine, achieving better accuracy. The use of a handwriting recognition engine with a higher accuracy than e.g. Calligrapher, in combination with our demonstrated interventions, may even improve the overall accuracy. Our methods on operating on the results of the handwriting recognition engine operate context independent. Using a dictionary to add the likelihood of upcoming characters may improve the accuracy in that part of the problem regarding confusable pairs, such as "r" and "v". Because of typing in characters one by one, a word completion feature could be added to handwriting recognition too. This also would increase the writing speed.

ACKNOWLEDGEMENTS

REFERENCES

Anantharaman, V. & Han, L. S. (2001) Hospital and emergency ambulance link: using IT to enhance emergency pre-hospital care. *International Journal of Medical Informatics,* 61, 2-3, 147-161.

Baumgart, D. C. (2005) Personal digital assistants in health care: experienced clinicians in the palm of your hand? *The Lancet,* 366, 9492, 1210-1222.

Bunke, H., Roth, M. & Schukattalamazzini, E. G. (1995) Off-Line Cursive Handwriting Recognition Using Hidden Markov-Models. *Pattern Recognition,* 28, 9, 1399-1413.

Chittaro, L., Zuliani, F. & Carchetti, E. (2007) Mobile Devices in Emergency Medical Services: User Evaluation of a PDA-based Interface for Ambulance Run Reporting. In: Löffler, J. & Klann, M. (Eds.) *Mobile Response*. Berlin, Heidelberg, New York, Springer, 19-28.

Citrin, W., Halbert, D., Hewitt, C., Meyrowitz, N. & Shneiderman, B. (1993) Potentials and limitations of pen-based computers. *Proceedings of the 1993 ACM conference on Computer science*. Indianapolis, Indiana, United States, ACM.

Frankish, C., Hull, R. & Morgan, P. (1995) Recognition accuracy and user acceptance of pen interfaces. *Conference on Human Factors in Computing Systems*. Denver, Colorado, United States, ACM Press/Addison-Wesley Publishing Co.

Gader, P. D., Keller, J. M., Krishnapuram, R., Chiang, J. H. & Mohamed, M. A. (1997) Neural and fuzzy methods in handwriting recognition. *Computer,* 30, 2, 79-86.

Haller, G., Haller, D. M., Courvoisier, D. S. & Lovis, C. (2009) Handheld vs. Laptop Computers for Electronic Data Collection in Clinical Research: A Crossover Randomized Trial. *Journal of the American Medical Informatics Association,* 16, 5, 651-659.

Holzinger, A. (2005) Usability Engineering for Software Developers. *Communications of the ACM,* 48, 1, 71-74.

Holzinger, A. & Errath, M. (2007) Mobile computer Web-application design in medicine: some research based guidelines. *Universal Access in the Information Society International Journal,* 6, 1, 31-41.

Holzinger, A., Geierhofer, R. & Searle, G. (2006) Biometrical Signatures in Practice: A challenge for improving Human-Computer Interaction in Clinical Workflows. In: Heinecke, A. M. & Paul, H. (Eds.) *Mensch & Computer: Mensch und Computer im Strukturwandel*. München, Oldenbourg, 339-347.

Holzinger, A., Hoeller, M., Bloice, M. & Urlesberger, B. (2008a). *Typical Problems with developing mobile applications for health care: Some lessons learned from developing user-centered mobile applications in a hospital environment*. International Conference on E-Business (ICE-B 2008), Porto (PT), IEEE, 235-240.

Holzinger, A., Höller, M., Schedlbauer, M. & Urlesberger, B. (2008b). *An Investigation of Finger versus Stylus Input in Medical Scenarios*. ITI 2008: 30th International Conference on Information Technology Interfaces, June, 23-26, 2008, Cavtat, Dubrovnik, IEEE, 433-438.

Holzman, T. G. (1999) Computer-human interface solutions for emergency medical care. *interactions,* 6, 3, 13-24.

Kjeldskov, J., Skov, M. B., Als, B. S. & Høegh, R. T. (2004) Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in Field. *Mobile HumanComputer Interactoin - MobileHCI2005*. Springer Berlin / Heidelberg, 529-535.

Klann, M., Malizia, A., Chittaro, L., Cuevas, I. A. & Levialdi, S. (2008) HCI for emergencies. *CHI '08 extended abstracts on Human factors in computing systems*. Florence, Italy, ACM.

Lewis, J. R. (1999). *Input Rates and User Preference for three small-screen input methods: Standard Keyboard, Predictive Keyboard and Handwriting*. Human Factors and Ergonomics Society.

MacKenzie, I. S. & Chang, L. (1999) A performance comparison of two handwriting recognizers. *Interacting with Computers,* 11, 3, 283-297.

MacKenzie, I. S., Nonneke, B., Riddersma, S., McQueen, C. & Meltz, M. (1994) Alphanumeric entry on pen-based computers. *International Journal of Human-Computer Studies,* 41, 5.

MacKenzie, I. S. & Soukoreff, R. W. (2002) Text Entry for Mobile Computing: Models and Methods, Theory and Practice. *HUMAN-COMPUTER-INTERACTION,* 17, 2, 147–198.

Marti, U. V. & Bunke, H. (2002) Using a statistical language model to improve the performance of an HMM-based cursive handwriting recognition systems. *Hidden Markov models: applications in computer vision*. World Scientific Publishing Co., Inc., 65-90.

Phatware (2002) Calligrapher SDK 6.0 Developer's Manual.

Plamondon, R. & Srihari, S. N. (2000) On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 22, 1, 63-84.

Plotz, T. & Fink, G. A. (2009) Markov models for offline handwriting recognition: a survey. *International Journal on Document Analysis and Recognition,* 12, 4, 269-298.

Shi, B. & Li, G. (2006) VLSI Neural Fuzzy Classifier for Handwriting recognition. Patent US 7,146,037.

Strenge, M. (2005) Konzepte und Toolkits zur Handschrifterkennung.

Tappert, C. C., Suen, C. Y. & Wakahara, T. (1990) The State of the Art in On-Line Handwriting Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 12, 8, 787-808.

Xue, H. H. & Govindaraju, V. (2006) Hidden Markov models combining discrete symbols and continuous attributes in handwriting recognition. *Ieee Transactions on Pattern Analysis and Machine Intelligence,* 28, 3, 458-462.

Yaeger, L. S., Webb, B. J. & Lyon, R. F. (1998) Combining neural networks and context-driven search for on-line, printed handwriting recognition in the Newton. *Neural Networks: Tricks of the Trade*. Berlin, Springer-Verlag Berlin, 275-298.

DEPENDABLE DISTRIBUTED TESTING
Can the Online Proctor be Reliably Computerized?

Ariel J. Frank

Department of Computer Science, Bar-Ilan University, Ramat-Gan, Israel
ariel@cs.biu.ac.il

Abstract: Distributed Education (DE) enables education, both teaching and learning, anytime, anywhere, using any media or modality, at any pace. Assessment, especially testing, is a major component in the overall evaluation of an educational program. However, there still is an important component missing from most DE programs to enable its full realization in a distributed environment: dependable distributed testing (DDT). The paper presents a comprehensive risk analysis of dependable distributed testing that classifies seven (types of) risks, introduces the resultant DoDoT (Dependable observable Distributed online Testing) reference model, and examines its coverage by three commercial systems. However, these systems are not yet in use in most DE frameworks and do not have yet full DoDoT coverage. The vision of the DoDoT reference model is the continued pursuit and adaptation of new, innovative technologies and methods to make dependable distributed testing increasingly more computerized, reliable, affordable and prevalent.

1 INTRODUCTION

The incessant evolution of the Web has also accelerated the adaptation of Distributed Education (DE), a generalization of Distance Learning, which is in wide use nowadays for educational aims (Allen & Seaman, 2010). DE enables education, both teaching and learning, anytime, anywhere, using any media or modality, at any pace.

There are many important topics in education in general and in distributed education in particular (Simonson, Smaldino, Albright & Zvacek, 2009). We concentrate here on the topic of assessment – of learners by teachers – a major component in the overall evaluation of an educational program. Assessment, be it summative or formative, is a process dealing with control, measurement and the systematic documenting of learners' achievements so as to be able to examine to what extent have the learners advanced towards predefined educational goals. Learners' assessment can be achieved using some combination of tools, such as assignments, papers, projects, quizzes, open and closed tests, forums, group work, etc.

The focus here is on the classic assessment tool – the test (exam, quiz), taken by a testee (examinee, test taker), usually a student, pupil, or trainee, learning in an organization associated with higher education, high schools, or business, respectively. But how can testing security be provided to assure testing integrity? Obviously, there are people who will cheat or defraud if the stakes are high enough and the deterrence too low (Bailie & Jortberg, 2009; King, Guyette & Piotrowski, 2009; McCabe, Trevino & Butterfield, 2001; VS1, n.d.).

Tests can be paper-based or computer-based. Modern educational frameworks (e-learning) provide for computer-administered testing (e-testing). However, notwithstanding the significant growth of DE and its recent ubiquity, there still is an important component missing from most DE programs to enable its full realization in a distributed environment: dependable distributed testing (DDT) (Bailie & Jortberg, 2009; Graf, 2002; Guess, 2008). The term "dependable" here means that distance testing is available to all testees in a secure, reliable manner that maintains testing integrity. The focus here is only on the technical aspects of DDT.

The contribution of this paper is in the comprehensive risk analysis of DDT, in the resultant DoDoT (Dependable observable Distributed online Testing) reference model and in examining its coverage by three commercial systems. The rest of the paper is structured as follows. The following section discusses some relevant aspects of dependable testing. In Section 3, seven (types of)

risks in DDT are analyzed. In Section 4 the resultant DoDoT reference model is introduced and its coverage by three commercial systems is examined in Section 5. We conclude the paper with Section 6.

2 DEPENDABLE TESTING

Before delving into the risk analysis of DDT, the following subsections review some relevant aspects of dependable testing: accreditation of DE programs, in-person proctoring, and testing integrity.

2.1 DE Program Accreditation

Are DE programs as reputable as regular educational programs? Or more formally, are DE programs regulated and accredited as non-DE programs are? Specifically, accreditation requires that an organization offering an educational program must positively prove by some documentation that an enrolled person is the same learner who does the work leading to graduation (Acxiom, n.d.; VS2, n.d.).

Admittedly, there is a trend in DE not to rely on high-stakes testing for learners' assessment (Foster, 2008). This trend advocates worry that they will be forced to use a particular assessment process that could turn out be too expensive or that would overemphasize tests. Their driving idea is that DE teachers should rely more for assessment on written assignments, threaded discussions, blogs, wikis, e-portfolios, online quizzes, open tests and capstone projects. The assumption is that teachers can become familiar with their learners' working and writing styles in order to spot cheating and fraudulent work, and in time be able to individually assess them.

However, educational organizations requiring high-stakes assessment probably need to depend on testing (Heubert & Hauser, 1999). The lack of dependable testing might turn out to be a major obstacle in complying with required accreditation. Having such "validated testing" could garner more respect for DE programs and raise their credibility with regulators. However, even if distance testing is provided, it is recommended that learners should not be forced to take an online test. In the same way that usually there is a choice if to take an educational program on-ground or online, so should there be a choice for test taking.

2.2 In-person Proctored Tests

Dependable testing is usually realized by human oversight, in the form of proctors who administer the test. Dependable test delivery, with in-person proctors, can be realized using one of three options (Questionmark, n.d.):
1. On-site testing
2. Off-site testing centers
3. One-on-one proctoring.

On-site testing necessitates the physical arrival of all proctors and testees (and usually their teachers) to the organization at the same time. With computer-administered testing, each testee works on a computer that is organizationally preloaded with the test and other tools authorized for use. This is a major undertaking by all involved, though it does provide dependable testing.

Off-site testing centers necessitate the physical arrival of proctors and testees to the closest testing center at similar times (time zone adjusted). This requires travel of staff and faculty to proctor tests taken at off-site public facilities or at third-party testing centers. With computer-administered testing, the organizational logistics required to support multiple testing centers in parallel is a complex effort. It can turn out to be a cumbersome operation, though it also provides dependable testing.

One-on-one proctoring necessitates the recruitment of an in-room proctor to supervise a certain testee at a pre-designated place and time. That is, the testee usually has to make all arrangements for selecting a certified proctor (former teacher, work colleague, librarian, etc.), which will physically oversee the taking of a secured test delivered on time to the testee computer. Overall, this is a burdensome option with a lower testing dependability, since there is higher opportunity for collusion between a not-so-reputable testee and a not-so-reputable proctor.

As aforementioned, each of these in-person proctoring options requires much planning and execution, involving relevant DE administrators, proctors, learners, and teachers, and the appropriate computer-administered testing infrastructure, to solve what can turn out to be "a logistical nightmare". Any of these options, above and over the, often repeated, efforts and costs, can also lead to growing frustration among the distance learners since these options contradict the core principles of DE and its full realization.

2.3 Test Integrity

The prevalent assumption is that traditional proctors and secure logins suffice to ensure honest computer-administered testing (Carter, 2009). However, any (testing) security system can be circumvented if the will is there and the capabilities exist. In-person proctored tests are also not foolproof. Proctors never

do anything but keep honest testees or uncreative ones honest. How does one make sure that no one cheats in an auditorium attended by a large number of testees? Or as another example, consider a testee with a hidden pinhole camera on a shirt button that broadcasts the screen content to an accomplice who uses a cellphone to advise the testee that wears a miniature wireless earphone concealed by long hair. What should be possible is to inhibit testing integrity risks, not fully prevent them (Graf, 2002; Weippl, 2005).

3 ANALYSIS OF DDT RISKS

Assuming no in-person proctors in DE frameworks makes dependable distributed testing even tougher to realize, but does not turn it into "mission impossible". The fact that (at least) three commercial DDT systems are in use is an indication of a need for such DE systems (Section 5). As part of a literature survey (see References) and extensive research leading to the proposed DoDoT reference model (Section 4), we conducted a comprehensive risk analysis of the cheating and fraud options attemptable during tests. Following, we analyze seven (types of) risks that were identified and classified: testing mismanagement, impersonator, PC misuse, forbidden stuff, accomplice, test leakage, and "electronic warfare".

3.1 Testing Mismanagement

Dependable distributed testing requires varied services and tools for test creation, test data management and archiving, test scheduling, test delivery, test grading, test reporting and test security. The potential for testing mismanagement is huge. Fortunately, the following coupled DE systems can handle this risk.

DE frameworks are usually based on a Learning (Content) Management System (LMS/LCMS) – a software application, usually web-based, for the administration, documentation, tracking, and reporting of e-learning programs, classroom and online events, and educational content (Ellis, 2009). The companion system, Testing Management System (TMS), has similar functionalities but specializes on the testing domain. A TMS can be integrated into an LMS or be interconnected with existing one.

The TMS simplifies the entire test management process while adhering to privacy policies. It should provide testing security for the entire lifecycle of a test. Teachers can post tests from anywhere and learners can take them anywhere with computerized test delivery and full documentation of the testing process. The TMS provides teachers with complete control over and access to the test items, test results and other test information. It can track testees' progress throughout the test sessions, reducing the tension and administrative burdens normally associated with the test season. A TMS can also be used to manage other assessments including unproctored practice tests, quizzes and take-home tests, but this is not the emphasis here.

To support DDT, a TMS should be rigorously secured by use of leading-edge test security and communication technologies or even be run on a secure network. For example, it should use data encryption technologies for test information, secure protocols, firewall protection, etc.

In terms of distributed systems, such a TMS is a "dependable system" where the dependability concept covers important requirements such as availability, reliability, safety, maintainability, integrity, security and fault tolerance (Kopetz & Verissimo, 1993). In a similar sense, the term used here – "dependable distributed testing" – is a generalization of the often used term of "secure online testing".

As aforementioned, there are many components to a testing management system. The focus here is on the test delivery aspects of a TMS. Commercial TMSs include: Perception Secure (Questionmark, n.d.), Respondus (Respondus, n.d.). Section 5 reviews three commercial DDT systems that are TMS based: ProctorU (Pupilcity, n.d.), Online Proctoring (KryterionOLP, n.d.), Remote Proctor (Securexam, n.d.).

3.2 Impersonator

A serious DDT risk is an impersonator testee. How to verify that a learner, signed up for a DE program, is the same one taking the test if the testee is far away? The identity check required when taking a test can be realized by testee verification (Acxiom, n.d.; Bailie & Jortberg, 2009; Schaefer, Barta & Pavone, 2009; Weippl, 2005). Baseline verification of testee identity in computer-administered testing can be achieved by authenticating the username and password during testee login.

The problem though is how secure are usernames and passwords? Learners employing someone else to take their test instead of them would willingly share their username and password with the impersonator, regardless of any rules or regulations. Similarly, we cannot rely on common information about testees (e.g., identification number, mailing address) or even something supposedly only testees know but

that is a "shared secret" with the TMS (e.g., mother's maiden name, favorite color). The problem is that "what you know" is easily sharable. We also cannot rely on some artefact the testees have (e.g., driver's license, smartcard, wearable RFID tag), i.e., on "what you have". Note that we should also not rely on the location of the testees (e.g., IP address, GPS tracking device), i.e., "where you are", so as not to limit the testees in where they take the test.

So how can impersonation be prevented in DE environments? One solution is to achieve testee verification using biometric enrollment and authentication processes (i.e., "what you are"). There are several biometric verification technologies that could be considered (Prabhakar, Pankanti & Jain, 2003). Some are already in use in DE frameworks: fingerprint verification (Securexam, n.d.), face verification (Kryterion, n.d.; Securexam, n.d.), signature verification (StudentPen, n.d.). As part of the authentication process, there is a need to decide when (at start, periodic or random) and how to authenticate the testee and what are the consequences of failure to authenticate. (These processes are also important to assure non-repudiation of test taking.)

Some researchers have coined the term "behaviometrics" for behavioral biometrics such as typing rhythms patterns or mouse movements (i.e., "what you do"), where this analysis can be done continuously without interrupting or interfering with user activities. For example, Webassessor (Kryterion, n.d.) uses keystroke analysis for recognizing unique typing styles. It measures the pattern of keystroke rhythms of a user and develops a unique biometric template of the user's typing pattern for future authentication.

Another testee verification option is use of the challenge questions methodology to inquire on personal history that only the testee can answer for (i.e., "what only you know"). For example, Acxiom Student Identity (Acxiom, n.d.; Bailie & Jortberg, 2009) poses in real-time a few targeted questions that challenge the testee and scores the answers. Challenge questions can be based on third-party data retrieved from large-scale public or private databases, while maintaining privacy policies. Strategies can be used to determine which challenge questions to ask, how many questions, passing thresholds, and red flags on fraud indicators. For example, challenge questions could be asked at sign-on, periodically or also at random. Unlike biometric and behaviometric authentication, the challenge questions methodology does not require pre-test enrollment.

3.3 PC Misuse

Nowadays, most learners have their own personal computer (PC) or can easily gain access to one for testing purposes. We assume the testee uses a well-equipped PC. For purposes like voice verification, environment sounding, or audio chat, the PC requires a microphone. In addition, speakers need be connected to the PC unless the testee wears headphones. For purposes like face verification, environment watching, or video chat, the PC requires a (preferably sound-equipped) webcam. Moreover, for DE purposes, PCs need broadband connections to access the Internet from anywhere, anytime. We do not relate here to mobile learning (m-learning) and its varied devices and connections.

With computer-administered open tests, the testee can access local files, use PC applications or browse the Web. Web browsers are usually designed to be as open and flexible as possible. The focus here though is on high-stakes closed tests. Consequently, when delivering such tests, there is need for more security than is available on a regular PC or that a common Web browser provides (Questionmark, n.d.).

However, with no in-person proctors, how can the entire test session be secured to ensure the testing integrity? For example, the testee could have installed some software on the PC before the test for the express purpose of defeating the test security. Or as another example, there is always a temptation to Google for help. The solution is to use PC lockdown software to secure the testing environment and its test contents. However, how can the lockdown software be securely activated on the PC to ensure its continued reliable operation? The solution is to securely access and activate the PC lockdown software via the TMS and keep them interoperating.

With lockdown software running on the PC, it can be ensured that the test is only delivered via the organization's TMS, after successful biometric enrollment and authentication processes. PC lockdown software includes varied tools that enable lockdown of the desktop, operating system, and Web browser. The idea is to flexibly restrict or completely disable testees' access to the compromising functionalities of these resources. Besides access to the test questions and use of authorized files or tools (e.g., word processing, spreadsheet analysis), the lockdown software secures the testing PC by preventing print, capture, copy, or access to other locally stored or Web accessible files and programs.

PC Lockdown software usually disables (if not restricts) the following functionalities:

- Cut/copy/paste of data to/from the testing environment

- Screen capture/printing functions
- Control/function keys/shortcuts
- Task/application start/access/switch
- Right-click menu options
- Menu options or icons activation
- Setting of PC date/time
- Pop-up windows
- Messaging, screen sharing, network monitoring.

In addition, browser lockdown usually disables (if not restricts) the following functionalities:

- Search/surf the Web
- Browser menu and toolbar options with possible exception for Back/Forward/Refresh/Stop
- HTML source code viewing
- Cache/store of pages in history/search listings. Moreover, PC lockdown software can provide for the following requirements:
- Automatically start at sign-on page of the organization's TMS.
- Testees cannot commence a test until they are provided with a special password by the TMS.
- The test questions are displayed in a full-screen mode that cannot be minimized.
- Following test completion, all test-related files are automatically submitted back to the TMS.
- Clearing of any cookies, caches, and temporary files at test session end.

When a test is launched, the testee is locked up into the testing environment (i.e., cannot suspend or exit it) until the test is submitted back. The testing environment should be able to withstand (un)intentional actions or breakdowns, shutdowns or restarts, and network disconnections, and be able to recover the testing environment and contents. With advanced technologies such as "software as a service", Web services, virtualization and cloud computing, such robust testing environments can be nowadays supported.

As another option, "Remote Desktop" software can be used to observe the testee screen and even control the PC if deemed necessary by an online proctor. It can also be used to assist the testee and provide technical support if need be. The testee must have given previous authority for remote access to the online proctor.

Commercial PC lockdown software include: Simpliciti (Simpliciti, n.d.), KioWare Lite (KioWare, n.d.), Perception Secure (Questionmark, n.d.), Respondus (Respondus, n.d.). There is also an open-source Safe Exam Browser (SEB, n.d.).

3.4 Forbidden Stuff

In regular closed tests, the testee puts away all forbidden stuff such as notes, reference sources, textbooks, computers, cellphones and other devices (King et al., 2009). But how can this restriction be enforced in the absence of in-person proctors?

The solution is online monitoring to proctor the testing environment to detect anything forbidden. Online monitoring, using real-time audio and video (A/V), can hear and watch the testees and their surrounding environment, while they enroll, authenticate and take the test on their PC. It can be carried out by (live or random) online proctor observation or by (continuous) test session recording that consists of A/V, biometrics and other testing event data that is sent to the TMS. Online proctors can observe the testee using one-on-one videoconferencing technologies. Test session recording uses streaming technologies where the A/V stream can also be viewed by an online proctor. Computerized processes can detect aberrances and use red flags to real-time alert the online proctor or indicate need for post-test analysis.

However, with a common webcam, there is a problem detecting forbidden material displayed at the room back or hidden behind or below the PC. We provide a solution for this in the next subsection.

3.5 Accomplice

Another serious DDT risk is a testee accomplice (Eplion & Keefe, 2007). How can an accomplice be prevented from aiding the testee? There is a need to disallow the same accomplice means used in a regular test such as exchange of notes, use of cellphones, rendezvous at the toilets, etc. The distance testee should be required to disconnect all phones, not leave the room, not let another person enter the room, etc.

Online monitoring can also be used to detect an accomplice via a sound-equipped webcam. However, a regular webcam isn't enough to ensure testing integrity. For example, a video projector in back of the room or a hidden (pinhole) camera in front can project the screen content to an in-room accomplice standing behind the PC. The accomplice can in return signal the testee (say for multiple choice questions), use sign language, or write answers onto a raised (hand-held) whiteboard.

Asking the testee to physically pan the webcam around the PC to check on the surroundings is an awkward process, especially if it has to be repeated during the test itself. A better solution is to use a 360° webcam. For example, the Securexam Remote Proctor (SRP) unit (SecureExam, n.d.) encloses a 360° webcam. The unit features a mirrored sphere suspended above a small pedestal. The sphere reflects a deep 360° view around the testee, which

the webcam picks up. A 360° webcam can be used to detect an accomplice, as well as use of forbidden stuff, also behind and below the PC, and red flag online monitoring that something might be awry. It is hard to cheat without some suspicious sound or motion being made by the testee or the accomplice.

Another countermeasure is to detect and obstruct any (hidden pinhole) camera by use of an inexpensive, simple laser pointer – not damaging to humans – to zap (blind) the camera, thereby generating a camera capture resistant environment (Naimark, 2002). Similarly, a long video cable or hidden camera can transmit the screen content to an off-site accomplice who uses a cellphone to advise the testee who wears a miniature wireless earphone. The countermeasure is the use of cellular detectors and jammers (Wollenhaupt, 2005).

Advanced recognition technologies could also be put to use. For example, if the testee decides to play music to relax while taking the test, voice/sound recognition can disregard it. As another example, image/object recognition can prevent a false alarm if a pet suddenly wanders around the room or jumps on the testee lap.

3.6 Test Leakage

An acute DDT risk is test leakage, especially for same time tests (Eplion & Keefe, 2007). Although testees can (try to) memorize (some of) the closed test's content, at least they should not be able to compromise it at test time. The use of a secure TMS and PC lockdown software prevents many of the options for test leakage. Restrictions enforced for the accomplice risk also apply. Options to prevent test leakage via an accomplice have also been covered.

However, in regular closed tests, scrawling is a natural test activity that is allowed in the test book ("blue book" in USA). Similarly, scrawling in computer-administered testing can be allowed in a digital notebook ("private workspace") that is part of the secured testing environment. Forbidden writing to paper can be detected by online monitoring. However, indirect recording of test questions by a testee that seemingly just reads the questions aloud is hard to detect (if the recording device is hidden), so such systematic reading aloud should be disallowed.

To hinder the leakage of a test, its questions and answers, one or more of the following or similar methods, collectively named here "Schemed questioning", could be considered:

- Use of test banks with random selection of questions.

- Scrambling the order of questions and answers (for multiple choice questions).
- Presenting just one question at a time.
- Setting time allotments for question answering (timed test delivery).

3.7 "Electronic Warfare"

The concern here is with the physical protection of the PC hardware and devices. How can the PC and especially its devices such as the camera, microphone and biometric devices be protected from tampering? There is a need to detect lost A/V signals or loss of feed quality. As another problem, the webcam real-time A/V stream could be substituted by a pre-recorded one.

The detection of this can be done by online monitoring. However, to discourage more advanced "electronic warfare", i.e., disabling or circumventing the capability of these devices, a separate hardware proctoring unit that physically encloses the devices can be used. To be easy to use, the unit should be portable and pluggable, say via USB. The proctoring unit has to be of course first acquired by the learner before any testing activity. To ensure the testing integrity, as part of an enrollment process, the unit should be remotely registered to both the testee and the PC used for test taking. The proctoring unit itself should be physically tamperproof, and secured by the TMS and PC software lockdown so as to red flag any mishandling of the unit.

4 DoDoT REFERENCE MODEL

Based on the above DDT risk analysis, we introduce the resultant DoDoT/RM (Dependable observable Distributed online Testing Reference Model). (Dodot stands for aunts in Hebrew – it is slang for the traditional elderly female proctors.) DoDoT/RM suggests an array of specific methods that can answer the seven risks (see mapping in Table 1), so as to enable the reliable computerization of online proctoring in DDT systems. The paper's author is unaware of any published similar attempt to define a dependable distributed testing reference model.

The premises of DoDoT/RM are as follows. To assure DDT, each and every one the seven (types of) risks should be covered. Moreover, not just one, but at least two of the proposed methods should be used for risk mitigation. The idea is to make cheating and fraud significantly hard – too expensive for the testee to make it worthwhile taking the risks. For example, for testee verification it is recommended to

use two-factor authentication, where two different factors (out of biometrics, behaviometrics and challenge questions) are used in conjunction to deliver a higher level of authentication. Similarly, both online proctor observation and test session recording can be used for more reliable monitoring. However, since most of the suggested methods can concurrently answer several risks, just a minimal covering set of methods should be chosen.

The monitoring of the test session should be online, not just offline, since the test environment should be real-time observable and the testee be made aware of it. However, online monitoring does not necessarily require that a human proctor continuously observe the testee. Since a human proctor is an expensive resource, live observation could be done at test launch, randomly, or if real-time computerized red flags were raised. It also does not have to be achieved via a one-on-one videoconference. The A/V stream of the test session recording received at the TMS can be observed by an online proctor. Note also that there is no, or less, need to repeat authentication processes if there is test session recording that can be post-test analyzed.

However, how many proctors can be employed concurrently? And of those, how many are capable to diligently watch and listen to hours of testees A/V during or after the test taking with the possibility in mind that testees might cheat at some point? Considering the human limitations in continuous monitoring, the premise of DoDoT/RM is that computerized processes are preferable in this regard to human ones. Most, if not all, of the online proctor observations can be replaced by computerized processes (possibly adaptive AI agents) that can deter and detect aberrant events and red flag them as real-time alerts for the online proctor or as signals indicating a need for post-test analysis.

The online proctor could then monitor a testee just a few times during a test to observe if what is being done matches the sounds and actions on the testee PC. If there are red flags, the test session recording can be later analyzed, preferably again by computerized processes, which provides in addition the recorded proof of any wrongdoing.

To inhibit the risks of accomplice and test leakage, it is recommended to use technologies for hidden devices obstruction such as camera zapping and cellphone jammers. For sophisticated computerized monitoring and post-test analysis, use can be made of advanced recognition technologies such as voice/sound and image/object recognition. Post-test analysis of multiple tests can be used to detect aberrant trends, for example, by data mining.

As aforementioned (Section 3.5), use of a common webcam is not enough – only a 360° webcam provides continuous view of the entire testee surroundings. Moreover, having separate PC devices such as 360° webcam, biometric devices, camera zappers and cellphone jammers is problematic (Section 3.7). Use of a separate proctoring unit to enclose all the devices is required.

A cost-effectiveness analysis of DoDoT/RM still needs to be carried out. However, it is assumed that a covering set of its methods, and specifically the separate proctoring unit, can be realized in a cost effective way (subsidized by the organization or priced at few hundred dollars overall per testee). For such and further technical considerations refer to (Acxiom, n.d.; Bailie & Jortberg, 2009; Foster, 2008; Jortberg, 2009). Clearly, when choosing such a covering set, other relevant aspects such as social, legal, ethical, economical, and psychological ones need also be considered (Schaefer, 2009).

5 COMMERCIAL DDT SYSTEMS

For a feasibility check of DoDoT/RM we examine its coverage by three commercial DDT systems (Bailie & Jortberg, 2009; Foster, 2008): Pupilcity ProctorU, Kryterion Online Proctoring, Securexam Remote Proctor. Due to paper space constraints, the review focus here is on their outstanding techniques and services (more detailed information is on their websites). Note also that no caught cheating rates are made public by these companies. For each system, we mark in Table 1 the methods that are in use by a checkmark and those not in use by a dimmed x. Subsection 5.4 compares these working DDT systems regarding their DoDoT/RM coverage.

5.1 Pupilcity ProctorU

ProctorU (Pupilcity, n.d.) allows learners to securely take tests online by using videoconferencing to connect one-on-one with live, certified proctors and follow their instructions. ProctorU was originally developed for internal use at Andrew Jackson University (AJU) and was later spun off into a separate company, Pupilcity (Morgan, 2008). ProctorU uses the Acxiom Identify-X technology for a real-time online identity verification service that uses the challenge questions methodology (Acxiom, n.d.). This technology was piloted and put to test several times at National American University (ANU) (Bailie & Jortberg, 2009). ProctorU is affiliated with 20 educational institutions.

Table 1: DDT risks and methods.

Risks	Methods that can answer the seven risks	Pupilcity ProctorU	Kryterion OLP	Software Secure SRP
Testing Mismanagement	• Testing Management System	✓	✓	✓
	• Secure communication	✓	✓	✓
Impersonator	• Biometric authentication	×	✓	✓
	• Behaviometric authentication	×	✓	×
	• Challenge questions	✓	×	×
	• Online proctor observation	✓	✓	✓
	• Test session recording	×	✓	✓
PC Misuse	• PC Lockdown Software	×	✓	✓
	• Remote Desktop Software	✓	×	×
	• Aberrance computerized red flags	×	✓	✓
Forbidden Stuff	• Online proctor observation	✓	✓	✓
	• Test session recording	×	✓	✓
	• Aberrance computerized red flags	×	✓	✓
Accomplice	• Online proctor observation	✓	✓	✓
	• Test session recording	×	✓	✓
	• Use of 360° webcam	×	×	✓
	• Hidden devices obstruction	×	×	×
	• Advanced recognition technologies	×	×	×
Test Leakage	• PC Lockdown Software	×	✓	✓
	• Online proctor observation	✓	✓	✓
	• Test session recording	×	✓	✓
	• Use of 360° webcam	×	×	✓
	• Aberrance computerized red flags	×	✓	✓
	• "Schemed questioning"	×	✓	×
"Electronic Warfare"	• Detect lost A/V signal & feed quality	✓	✓	✓
	• Separate proctoring unit	×	×	✓

Each testee first needs to individually schedule a test at ProctorU's Online Proctoring Center. There are four steps in taking the test:

1. Connect – ProctorU automatically connects the testee to the online proctor.
2. Observe – proctor connects to testee's screen.
3. Prove identity – proctor watches the testee as he/she authenticates identity.
4. Monitor – proctor observes testee taking test.

ProctorU uses the TokBox video chat and Remote Desktop software. The online proctor watches the testee via a webcam as he types away at the keyboard, observes the screen, and listens for other sounds in the testing environment. Aberrant actions can be manually documented in the form of screen captures and camera shots that are sent to the TMS. ProctorU does not make use of the other methods suggested by DoDoT/RM.

5.2 Kryterion Online Proctoring

Webassessor (Kryterion, n.d.) is a secured online testing platform that provides a wide variety of testing technologies and services. Kryterion introduced Webassessor's Online Proctoring (OLP) system (KryterionOLP, n.d.) in 2007. A series of OLP pilots was carried out in conjunction with World Campus, the online arm of the Pennsylvania State University system (Shearer, Lehman, Hamaty & Mattoon, 2009).

It uses the Akamai secure network to provide robust testing delivery. OLP uses the Sentinel Secure technologies to lockdown the PC and conduct face verification and keystroke analysis for testee enrollment and authentication. The testing environment is continuously monitored and a testing session recording is generated. OLP utilizes varied security technologies and processes to deter and detect aberrance during the testing session and alerts online proctors when suspicious activities occur.

Kryterion employs certified, online proctors, called KCOPS, who can remotely observe and listen to as many as 50 testees at a time. They monitor a live video feed of each testee in real-time. Previous testing activity of testees is available to KCOPS for detecting aberrant behavior. Testee's aberrant behaviors or response time patterns (e.g., answering a question too fast or too slow) alert the KCOPS. OLP uses Real Time Data Forensics (RTDF) technology to red flag unusual events.

The KCOPS communicate with testees just via drop down menu options. KCOPS can send messages to the testee as necessary and take actions such as pausing, suspending or stopping the test based on testee behaviors and actions. For "Schemed questioning", testees receive questions one at a time after scrambling the order of test questions.

5.3 Securexam Remote Proctor

Software Secure (SoftwareSecure, n.d.) provides a suite of tools for secure online testing. The Securexam Remote Proctor (SRP) (SecureExam, n.d.) was an initiative of Troy University, which was commercially developed by Software Secure. It has been extensively experimented with and is long in use at Troy University (Powers, 2006; Guess, 2008). It was used for a pilot at a Small Southern Regional University (Bedford, Gregg & Clinton, 2009). SRP is affiliated with 15 educational institutions.

PlanetSSI is their web-based TMS. PC lockdown software is comprised of the Securexam Student and Securexam Browser. SRP uses biometric fingerprint verification and face verification, real-time A/V monitoring and recording of the testing session.

The SRP device is a separate proctoring unit that connects to the testee's PC as a USB plug-in. It includes a groove for scanning fingerprints, and a built-in 360° webcam. SRP interconnects with Securexam Browser and Securexam Student. SRP verifies the testee's identity through the use of finger-scan and face verification. Testees are recorded during tests and the recorded stream can be observed online. In addition, computerized filters can detect any suspicious changes in sound or motion, and red flag them for post-test analysis.

5.4 Discussion

Pupilcity ProctorU is a technically simple DDT approach since it mainly depends on online proctor observation and uses challenge questions for testee verification (Table 1). Consequently, it has only partial coverage of DoDoT/RM. ProctorU is more oriented to individual test taking than to same time testing. It does not have two-factor authentication since there is no biometric/ and behaviometric authentication; it relies solely on challenge questions. It does not use PC lockdown software, do test session recording or provide aberrance computerized red flags. It also does not have a separate proctoring unit having a 360° webcam. Kryterion Online Proctoring is a technically rich DDT approach with good coverage of DoDoT/RM (Table 1). OLP has two-factor authentication: biometric (face verification) and behaviometrics (keystroke analysis). It is noteworthy that this chosen two-factor authentication scheme requires no biometric device. OLP supports both online proctor observation and test session recording so any required balance between them can be realized. It has a varied set of computerized processes to real-time red flag aberrant actions and behaviors. However, it does not make use of a separate proctoring unit having a 360° webcam.

SRP has excellent coverage of DoDoT/RM since it also uses a separate proctoring unit with a 360° webcam (Table 1). However, it has only two-factor biometric authentication: face and fingerprint verification. It is noteworthy that SRP emphasizes test session recording while relying less on online proctor observation. It uses computerized processes to red flag suspicious activities by recording A/V clips for post-test analysis and aberrance proof.

Note that these systems could use more advanced recognition technologies for sophisticated computerization of processes to red flag aberrances. They could also utilize hidden devices obstruction (Section 4) to inhibit the associated risks. In any case, a DDT system with full or fuller coverage of DoDoT/RM has yet to be developed and deployed.

6 CONCLUSIONS

To increase the testing integrity of DE programs, there is growing need to deliver DDT anytime, anywhere. Wide deployment of DDT systems to achieve testing integrity has long been overdue. The introduced DoDoT/RM is based on a comprehensive DDT risk analysis. The fact that three commercial DDT systems are in use is an indication for their need. Nowadays, due to technological advances and improved methods, DDT systems can securely deliver high-stakes tests worldwide. These DDT systems utilize varied test security methods to deter and detect cheating and fraud by testees.

However, DDT systems are not yet in use in most DE frameworks. Moreover, these systems do not yet provide full DoDoT/RM coverage to enable reliable computerization of the online proctor – more experimentation and comprehensive field use is still needed. The vision of DoDoT/RM is the continued pursuit and adaptation of new, innovative technologies and methods to make dependable distributed testing increasingly more computerized, reliable, affordable and prevalent.

ACKNOWLEDGEMENTS

REFERENCES

Acxiom (n.d.). *Identity Verification to Support Academic Integrity,* Acxiom White Papers. http://www.acxiom.com/StudentIdentity

Allen, I. E. & Seaman, J. (2010). *Learning on Demand: Online Education in the United States, 2009.* Newburyport, MA: The Sloan Consortium. http://www.sloanconsortium.org/publications/survey/pdf/learningondemand.pdf

Bailie, J. L. & Jortberg, M. A. (2009). Online Learner Authentication: Verifying the Identity of Online Users. *MERLOT JOLT*, 5(2), 197-207. http://jolt.merlot.org/vol5no2/bailie_0609.pdf

Bedford, W., Gregg, J. & Clinton, S. (2009). Implementing Technology to Prevent Online Cheating: A Case Study at a Small Southern Regional University (SSRU). *MERLOT JOLT*, 5(2), 230-238. http://jolt.merlot.org/vol5no2/gregg_0609.pdf

Carter, D. (2009, July 10). ED OKs Proctors, Secure Logins for Online Tests. *eSchool News*. http://www.eschoolnews.com/2009/07/10/ed-oks-proctors-secure-logins-for-online-tests-2

Ellis, R. K. (2009). Field Guide to Learning Management Systems. *ASTD Learning Circuits*. http://www.astd.org/NR/rdonlyres/12ECDB99-3B91-403E-9B15-7E597444645D/23395/LMS_fieldguide_20091.pdf

Eplion, D. M. & Keefe, T. J. (2007). Practical Tips for Preventing Cheating on Online Exams. *Faculty Focus*. http://www.magnapubs.com/issues/magnapubs_ff/4_4/news/600136-1.html.

Foster, A. L. (2008, July 25). New Systems Keep a Close Eye on Online Students at Home. *The Chronicle of Higher Education, Information Technology*, 54(46). http://chronicle.com/article/New-Systems-Keep-a-Close-Eye/22559

Graf, F. (2002). Providing Security for eLearning. *Computers & Graphics*, 26(2), 355-365. doi:10.1016/S0097-8493(02)00062-6

Guess, A. (2008, October 8). From Blue Books to Secure Laptops. *Inside Higher Ed*. Retrieved Feb 3, 2010, http://www.insidehighered.com/news/2008/10/08/tests

Heubert, J. P. & Hauser, R. M. (Eds.). (1999). *High Stakes: Testing for Tracking, Promotion, and Graduation*. National Research Council, Washington, DC: The National Academies Press. http://www.nap.edu/openbook.php?record_id=6336

Jortberg, M. A. (2009). Student Authentication Solution Comparison. *Online Student Identity blog*. http://mikejortberg.blogspot.com/2009/06/student authentication-solution.html

King, C. G., Guyette, R.W. & Piotrowski, C. (2009). Online Exams and Cheating: An Empirical Analysis of Business Students' Views, *The Journal of Educators Online*, 6(1). Retrieved February 3, 2010, http://www.thejeo.com/Archives/Volume6Number1/Kingetalpaper.pdf

KioWare (n.d.). *KioWare Browser Lockdown Software*. www.kioware.com/default.aspx?source=kmshowcase

Kopetz, H. & Verissimo, P. (1993). Real Time and Dependability Concepts*,* In Mullender, S. (Ed.), *Distributed Systems*, Addison-Wesley, 411-446.

Kryterion (n.d.). *Webassessor*. www.kryteriononline.com KryterionOLP (n.d.). *Kryterion Online Proctoring*. http://www.kryteriononline.com/delivery_options

McCabe, D. L., Trevino., L. K. & Butterfield. K. D. (2001). Cheating in Academic Institutions: A Decade of Research. *Ethics & Behavior*, 11(3), 219-232. http://www.swarthmore.edu/NatSci/cpurrin1/plagiarism/docs/McCabe_et_al.pdf

Morgan, J. (2008). *Online Proctoring Perfected, Spun off by Tech-savvy University.* Retrieved February 3, 2010, www.webwire.com/ViewPressRel.asp?aId=80502

Naimark, M. (2002). *How to ZAP a Camera: Using Lasers to Temporarily Neutralize Camera Sensors*. http://www.naimark.net/projects/zap/howto.html

Powers, E. (2006, June 2). Proctor 2.0. *Inside Higher Ed*. www.insidehighered.com/news/2006/06/02/proctor

Prabhakar, S., Pankanti, S. & Jain A. K. (2003). Biometric Recognition: Security and Privacy Concerns. *IEEE Security & Privacy*, 1(2), 33-42. doi:10.1109/MSECP.2003.1193209

Pupilcity (n.d.). *Pupilcity ProctorU*. www.proctoru.com

Questionmark (n.d.). *Questionmark Perception Secure*. http://www.questionmark.com/us/perception/

Respondus (n.d). *LockdownBroswer*. www.respondus.com

Schaefer, T., Barta, M. & Pavone, T. (2009). Student Identity Verification and the Higher Education Opportunity Act: A Faculty Perspective. *Intl. J. of Instructional Tech. and Distance Learning*, 6(8), 51-58. http://itdl.org/Journal/Aug_09/article05.htm

SEB (n.d.). *SEB,* www.safeexambrowser.org/

Securexam (n.d.). *Securexam Remote Proctor*. http://www.remoteproctor.com/SERP

Shearer, R., Lehman, E., Hamaty, P. & Mattoon, N. (2009). Online Proctored Exams at a Distance: How can you Help Assure Academic Integrity?. *UCEA 94th Annual Conference*, Boston, Massachusetts.

Simonson, M. R., Smaldino, S. E., Albright, M. & Zvacek, S. (2009). *Teaching and Learning at a Distance: Foundations of Distance Education*. Allyn and Bacon.

Simpliciti (n.d.). *Lockdown Solutions*. www.simpliciti.biz

SoftwareSecure (n.d.). www.softwaresecure.com

StudentPen (n.d.). *Bio-Pen*. http://www.studentpen.com

VS1 (n.d.). *Integrity in Education Online*. http://www.virtualstudent.com/?page_id=19

VS2 (n.d.). *Accreditation in Online Degree Programs*. http://www.virtualstudent.com/?page_id=9

Weippl, E. R. (2005). *Security in E-Learning*. Advances in Information Security, Springer, Vol. 16.
Wollenhaupt, G. (2005). How Cell Phone Jammers Work, www.howstuffworks.com/cell-phone-jammer.htm

A GEOGRAPHIC INFORMATION SYSTEM (GIS) TO DEFINE INDICATORS FOR DEVELOPMENT AND PLANNING IN JORDAN

Balqies Sadoun[1] and Bassam Saleh[2]

[1]*Department of Architectural Engineering, Philadelphia University*
Jordan and on Sabbatical Leave from the Department of Surveying and Geomatics Engineering
Al-Balqa' Applied University, Salt, Jordan
[2]*Department of Surveying and Geomatics Engineering, Al-Balqa' Applied University, Salt, Jordan*
{balqiessadoun, bsaleh}@yahoo.com

Abstract: The computerization and the creation of a digital data base is a must for the employment of many of the state of the art tools in the Optimal Planning process. This is easy made using Geographic Information System (GIS). GIS is an effective modern planning technique which gives the power to create maps, integrate information, visualize and solve problems, present future ideas and develop valuable solutions in basically no time while relating information to a geographic component. The planning challenges today are mainly due to overpopulation, pollution, deforestation and immigration, which definitely have a geographic dimension. The Departments of General Statistics/Census in all countries are considered the main source of data for governmental and private agencies. In this work, we are using a Geographic Information System (GIS) to create a database system for Jordan (a developing country) to be used for optimal planning purposes. The created GIS for the country by governorates (12 Governorates) cover all aspects of life including social, economical, resources, among others. All the departments' data in addition to a large quantity of field data that we gathered are utilized to create GIS system. Results of many GIS analysis techniques are presented for demonstration purposes. Different kinds of data will allow concerned people to have proper planning, and development according to existing realities, and can aid in deciding on priorities of such plans. The digitizing of the data is a step will be a great step forward towards optimal and well informed decision making process in the country.

1 INTRODUCTION

Planning is a comprehensive tool to an enormous data to pick patterns, define relations and present the results to help in better and optimal decision making in virtually no time and cost. Planning the future of a country or planning everyday life (such as starting a new business, or finding the best soil for growing vegetables, or the best route to a location, etc.), has a geographical dimension and always related to a map. Thus, GIS role is vital such a process as it has the potential to offer the optimal solution.

GIS is a simulation methodology of all past, present and future situations at a minimal cost compared to other planning tools. It is a computer-based technology and methodology for collecting, managing, analyzing, modeling, and presenting geographic data for never-ending applications. It consists of a data base, map information and computer link to allow viewing, inquiring, interpreting, and visualizing data in many ways that reveal relationships, patterns, and trends in the form of maps, reports, and charts (Burrough, 1986), (Chrisman, 1999), (Ducker, 1979), (Star and Estes, 1990), (Andronache et al., 2006). Ducker (Ducker, 1979) defined GIS as the management of the data to retrieve new related data for ad hoc queries and analysis. Chrisman (Chrisman, 1999) considers GIS to play a role in the society as people measure and represent geographic phenomena, then transform these representations into other form while interacting with social structures. GIS changed the heart of planning and informed decision making

especially in the world of emergencies and life saving process. During Catharine storm in the USA, GIS (created for transportation purposes) was the only way to locate people during the hurricane when all other methods failed (Andronache et al., 2006). It helped in saving the lives of thousands of citizens and opened the eyes to the benefits of GIS technology. GIS and related technology will help analyze large datasets, allowing a better understanding of terrestrial processes and human activities to improve economic vitality, environmental quality, and response to emergencies as well as, finding the way to a supermarket. Today, GIS is a multibillion-dollar industry employing hundreds of thousands of people (planners, engineers, economists, etc.) and used extensively in all aspects of planning.

Jordan is a Middle Eastern country located in Southwest Asia. Administratively, it is divided into 12 governorates: Ajlun, Amman, Aqaba, Al-Balqa, Irbid, Jerash, Al-Karak, Ma'an, Madaba, Al-Mafraq, Tafilah, Az Zarqa (Fig.1). The department of General Statistics in Jordan categorizes its data according to governorates. GIS is used in this work to create a digital data to reflect the different nature of each governorate, its resources, population, etc. The created digital data for the country will allow all kinds of analysis by governorate to clarify and reflect their strong and weak points, needs. This will help in defining indicators for development and planning. ArcGIS software is used in our work for capturing, organizing, analyzing, mapping, and presenting spatial information.

Our main Objectives are to: (a) create a global digital data for the country by governorate, (b) allow the usage of endless capabilities of GIS analysis and (c) offer the possibilities of web posting and e-government utilization.
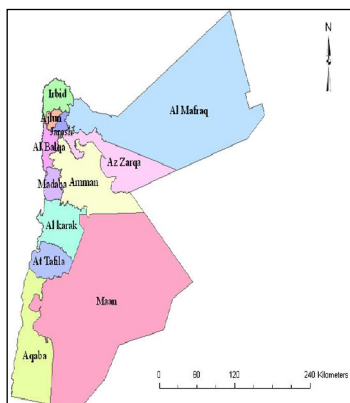
Figure 1: Jordan Governorates.

2 METHODOLOGY

GIS Data Model: GIS systems handle two types of data, raster and vector data. Raster data from scanned or remotely sensed images are poor at representing points, lines and areas, but good at surfaces. Vector data model uses points stored by their real coordinates, sequence of these points build lines and areas. The spatial data we used in building our GIS includes: Jordan Map (scale 1:1,000,000) and a Spot Satellite Image for Jordan. Global Positioning System (GPS) is used to collect the coordinates of important points to enrich our GIS such as: Universities (8 points), Hotels (58 points), Hospitals (29 points), Police Stations (44 points), Water Companies (11 points), Civil Status and Passport department (11 points), Municipalities (12 points), Airports (3 points). All points were taken according to Google Earth Map.

Attribute data give more information about features in tabular form. Attribute Statistical Data for Jordan for the years (1990 - 2005) were used to create the GIS database. All kinds of data such as: population count, population density, urban population, rural population, population by gender, number of births, deaths, marriages, and divorces, crimes, rain fall, water supply for domestic purposes, fuel distribution, number of new telephone subscribers, pharma- cies, registered engineers, post office mail boxes, post offices, road accidents, hotels, registered lawyers, charitable societies, bookshops, hospitals, health centers, dental clinics, schools etc. Available and field collected data were included in the work.
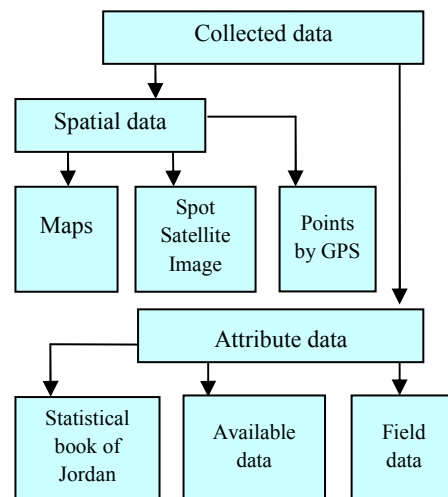
Figure 2: Collected data flowchart.

Data Processing: spatial and attribute data, maps and images were processed as follows: (a) scanning the map of Jordan to convert it to digital format, (b) defining the coordinate system; (Universal Transverse Mercator: UTM) and the Ellipsoid, (c) registering the scanned map (image) in the coordinate system using ground control points (points of known coordinates), and (d) creating the layers by digitizing the map (Fig.3).

Many layers were created using GIS. The output layers included: Jordan by Governorate map, cities, sea, water companies, Police stations, Municipalities, Hotels, airports, border stations, Civil status and passport departments, roads networks, universities, Governorate buildings, etc.

3 ANALYSIS AND RESULTS

GIS analysis is finding geographic patterns in the database and the relationships between features. The analysis methods can be very simple such as making a map (Fig.1) or more complex, involving models that imitate the reality of the situation, or by combining many data layers. The tabular data allows different analysis techniques such as classification, buffering and statistical analysis. The analysis achieved is in clear presentations and real interaction with the needed information.
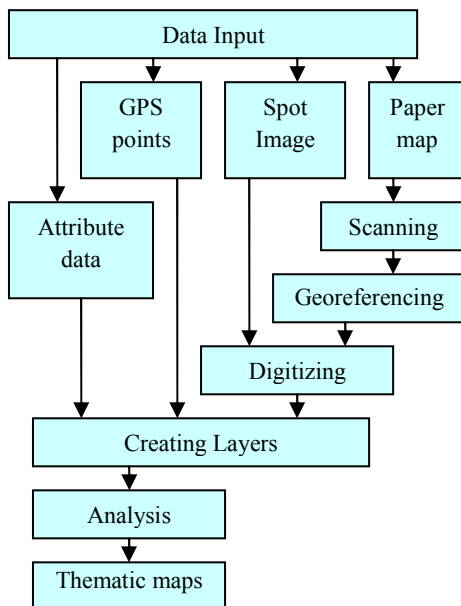
Figure 3: The Methodology.

Spatial Analysis: Queries offer a method of data retrieval from the data base, or on a new data

produced as a result of the data analysis. There are many methods of querying Data, which include: Identify, Find, Hyperlink, Query data by attribute, Query data by location.

Identify is the fastest tool for getting information about features by clicking on the feature, then all attribute data will appear in a tabular from. Find data is used to locate the position of such feature on the map in different layers upon the need, for example, to locate an airport or university on the map. Hyperlink data Hyperlink is a tool to obtain more information attached to features such as Photo or text (Fig.4).
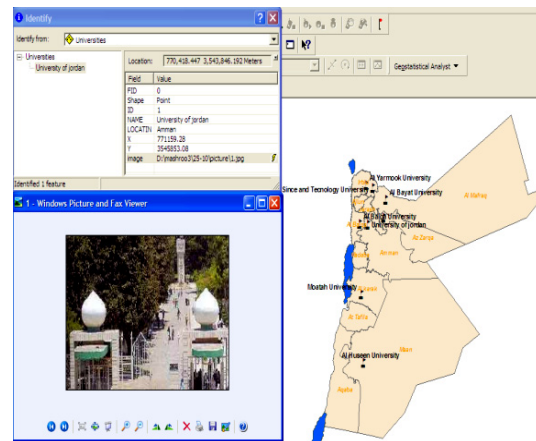
Figure 4: Hyperlink; the University of Jordan.

Query data by attribute. The features could be selected using the Standard Query Language (SQL), such as: Querying about the location of Universities in the northern city of Irbid (Fig.5).
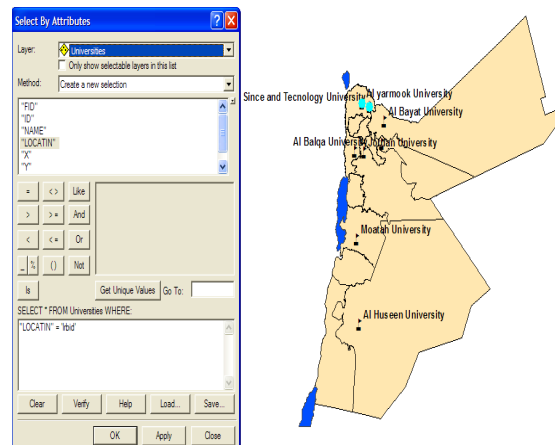
Figure 5: Universities in Irbid.

Figure 6 presents another type of queries such as
Querying about the Highways in Jordan which are
more than 75Km long.
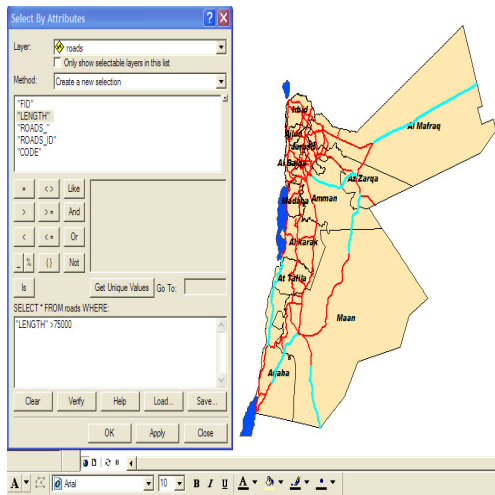
Figure 6: Highways more than 75Km long in Jordan.

Query data by location Selecting features by location
is a function that lets you select features from one or
more layers based on where they are located in
relation to the features in another layer such as: how
many "Police station" within (5km) from the Hotels
(Fig.7)? Hotels within (50km) from the Airports
(Fig.8)? and so on.
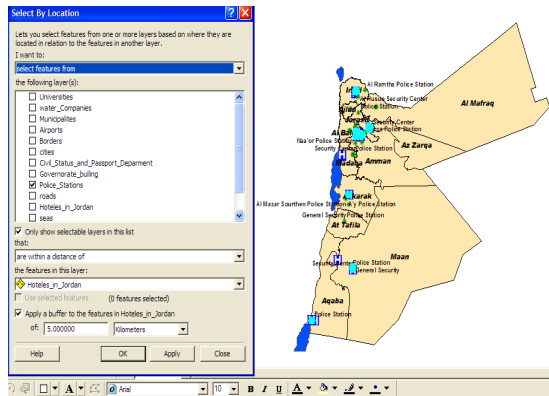
Figure 7: Police Stations within 5km from hotels.

Connectivity Analysis is done between points, lines,
and polygons in terms of distance such as: travel
time, optimum paths etc. Using the distance tool, we
can measure the distance between any two features
on the map. Figure 9 for example shows the distance
between Queen Alia airport and Aqaba airport.
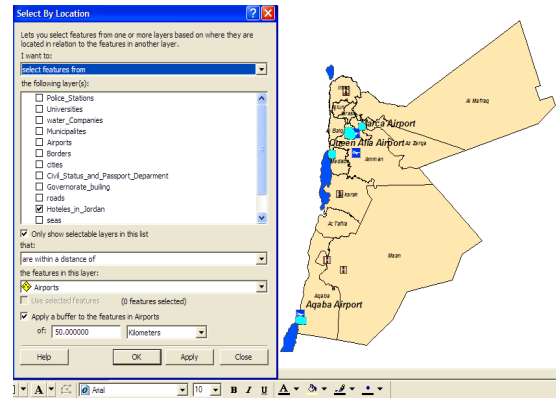
Figure 8: Hotels within 50 km from the Airports.
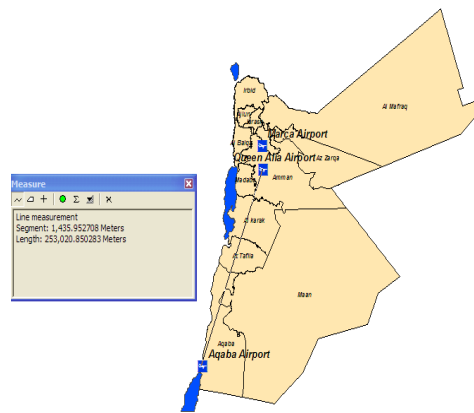
Figure 9: The distance between Queen Alia and Aqaba
Airports.
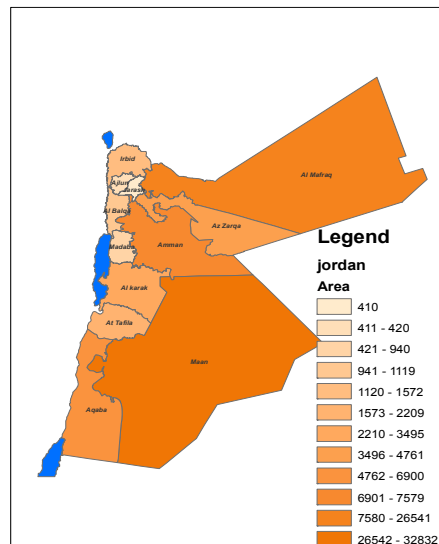
Figure 10: The Jordanian Governorates Areas in km2.

The Arc GIS software enables the conversion of a shape file into a feature class, which includes the area as an attribute file. Then, these areas could be used in statistical and spatial analysis. Figure 10 presents the areas of the governorates in km2 using classification technique.

Statistical Analysis entails the representation of the numerical data of the layers into graphical forms; a lot of mathematical functions are used to help in making engineering decisions.

Classification is used when we need to symbolize quantities, or want to see where attribute values lie in relation to one another on a continuous scale. Classification is an easy way for comparison in order to find or clarify any change in a situation or setting. It could reflect more than a result according to what is included in the classification process. Examples are many in city planning and in monitoring the results of the planning process in general. Figure 11 presents a comparison of the Population (number) for the years of 1990 & 2005 in a pie form. Figure 11 shows an increase in the population in all governorates. The pie plot is good only to show if there is an increase or decrease in each governorate, but it doesn't show the value of this increase in comparison between governorates as when using the histogram plot (Fig. 12).
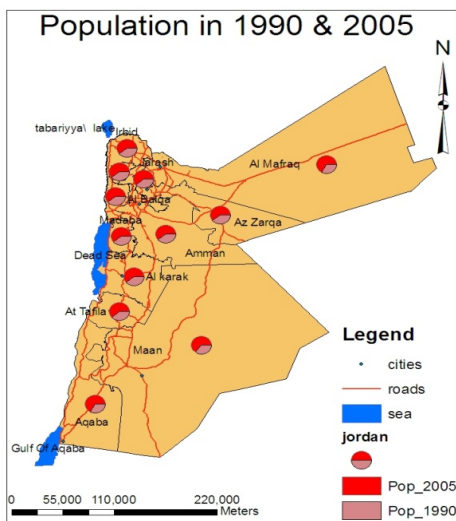
Figure 11: Population by governorate in the years 1990 & 2005.

Classification of population density for the years 2000 and 2005 (Fig.12) shows that Irbid Governorate has the most population density. Mean while, in Figure 13, we compare the population for the same years 2000 and 2005 and the results reflect

that Amman has the largest population and not Irbid governorate.
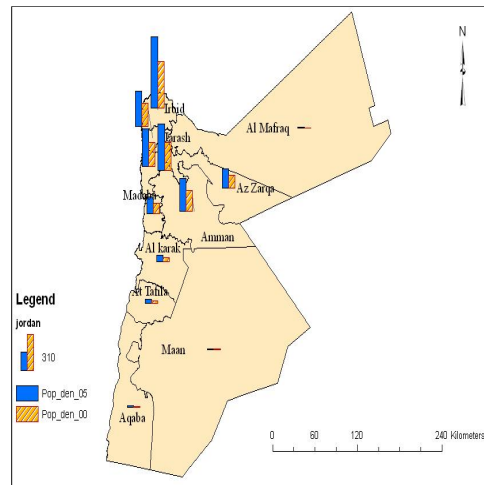
Figure 12: Population density by governorate in the years 2000 & 2005.

From a closer look to the population (number) by Rural and urban for the years 2005 and 2003, we can see clearly that the urban population in Jordan (blue color) is more important than the rural one.
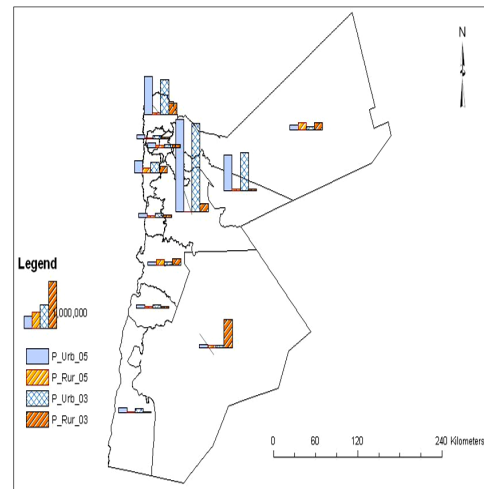
Figure 13: Population in 2005 & 2003 of the Urban and Rural.

Figure 13 shows that in year 2005 the rural population decreased alertly in the Maan governorate (2005). This may be explained as the incline in the agricultural area in the desert due to the lack of rain and water in general. In a previous study that we conducted using remote sensing, we found out the same results in Maan area. Another Classification

method of population density in 2005 could be used to clarify other relations.
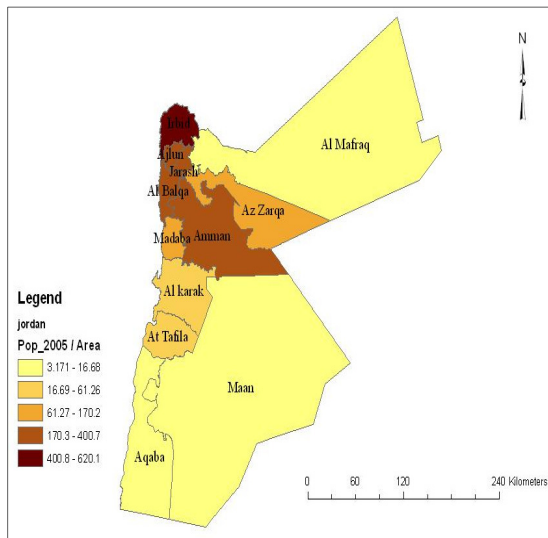
Figure 14: Population density (2005).

Figure 14 shows that the governorate of Irbid has the most population density in the country as it is the mainly agricultural governorate with the most rural population. The second in line is Amman, AlBalqa, Ajlun and Jerash as they all contain big cities and the least populated are the desert governorates with the least population density and resources. In this classification method we used the colors to join the governorates which have the same density. Another mode of presentation is the population density using dot representation (Fig.15).
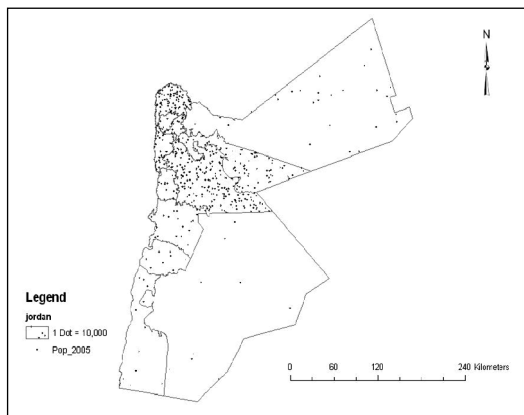
Figure 15: Population density in year 2005).
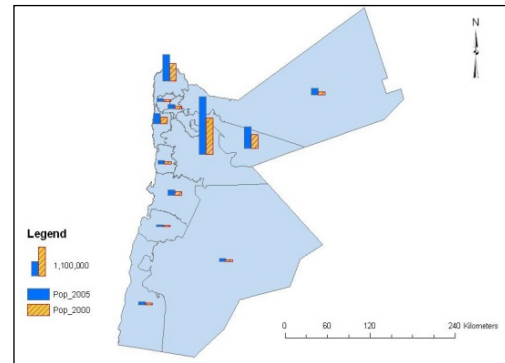
Figure 16: Population by gender in year 2005.

Figure 16 presents the Population in year 2005 by gender, 17 and 18 present other examples to be used in environmental (e.g, rain fall) and emergency analysis and planning. Endless possibilities and outcomes are offered using the created GIS system.
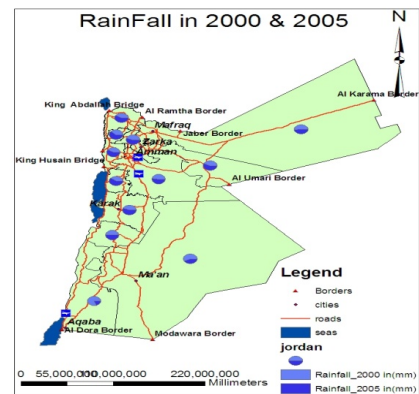
Figure 17: Rainfall in the years 2000 & 2005.

In Figure 18, we see that the highest General number of Crimes in 2005 is in the big cities. Again we can better understand reality of the statistics if we used percentages (crime/pop) or a histograms repre-sentation.
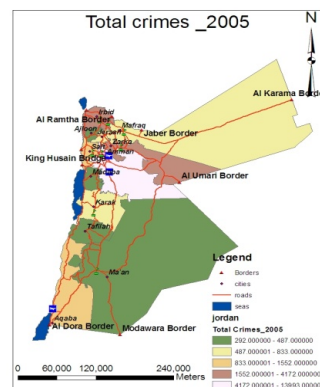
Figure 18: Numbers of General Crimes in year 2005.

4 CONCLUSIONS

To conclude, a GIS system is created for The Department of General Statistics in Jordan for all purposes especially in optimal and informed decision making and planning. The created digital database is easy to use. We can manipulate, maintain and update info effectively. In addition the data is presented in an easy to understand and act upon it. Finally, the ability of international information sharing through a web-site creation, collaborative environment and e-government development is possible with such a system.

REFERENCES

P. A. Burrough, "Principles of Geographical Information Systems for Land Resources Assessment". *Clarendon Press*, Oxford, 1986.

N. R. Chrisman, "What Does 'GIS' Mean?" Transactions in GIS, Vol. 3, No. 2, pp. 175-186, 1999.

K. J. Ducker, "Land Resource Information Systems: A Review of Fifteen Years Experience". Geo-Processing, Vol. 1, pp. 105-28, 1979.

J. Star and J. Estes, "Geographic Information Systems: An Introduction". *Prentice Hall*, Englewood Cliffs New Jersey, 1990.

C. Andronache, R. Hon, B. Mento, and R. Dalgin," "Mapping Hurricane Katrina with GIS", *Proceedings of the 2006 ESRI International User Conference Proceedings*, 2006. http://gis.esri.com/library/proc06/papers/papers/pap_2320.pdf.

SHORT PAPERS

E-BUSINESS APPLICATIONS IN ENGINEERING EDUCATION

Jeļena Zaščerinska

University of Latvia, Jurmalas Gatve 74/76, Riga, Latvia
knezna@inbox.lv

Andreas Ahrens

Hochschule Wismar, University of Technology, Business and Design, Philipp-Müller-Straße 14, 23966 Wismar, Germany
andreas.ahrens@hs-wismar.de

Keywords: e-Business applications, Engineering education, Needs analysis.

Abstract: Contemporary engineers need to become more cognizant and more responsive to the emerging needs of the market for engineering and technology services. Engineering education has the potential to contribute decisively to the sustainable development of e-business applications or enterprise 2.0, which penetrate our society more thoroughly with the availability of broadband services. However, the success of e-business applications in engineering education requires needs analysis to be considered. Aim of the following paper is to analyze the student engineers' needs in e-business applications of Web 2.0 technologies within engineering education on the pedagogical discourse. The meaning of the key concepts of e-business applications, engineering education and needs analysis is studied. The study shows a potential model for development, indicating how the steps of the process are related following a logical chain: determining e-business applications → revealing e-business applications in engineering education → defining needs analysis → empirical study within a multicultural environment. The results reveal that needs analysis of e-business applications in engineering education contributes to the incorporation of business-driven topics in the curriculum of engineering science.

1 INTRODUCTION

The primary target for software as a service is businesses (Vossen, 2009) considered as an overall attribute necessary for innovation-friendly societies to transform social capital into economic growth (European Union, 2008). New business ideas such as the payment service offered by RevolutionMoney, the mail service offered by eSnailer, the flight service offered by Virgin Charter, or the personalized TV service from Current.com are classical examples and have found widespread acceptance in the community (Vossen, 2009). E-business applications benefit from Web 2.0 where the increased data exchange within the system is no longer a limiting parameter with the current developments in the infrastructure. Aim of the following paper is to analyze the student engineers' needs in e-business applications of Web 2.0 technologies within engineering education on the pedagogical discourse. The meaning of the key concepts of e-business applications and needs analysis is studied. Moreover, the study demonstrates how the key concepts are related to the idea of engineering education

and shows a potential model for development, indicating how the steps of the process are related following a logical chain: determining e-business applications → revealing e-business applications in engineering education → defining needs analysis → empirical study within a multicultural environment. The remaining part of this paper is organized as follows: The introductory state-of-the-art demonstrates the authors' position on the topic of the research. Section 3 introduces e-business applications. E-business applications in engineering education are studied in section 4. The associated results of an empirical study are presented in section 5. Finally, some concluding remarks are provided in section 6 followed by a short outlook on interesting topics for further work.

2 STATE-OF-THE-ART

The modern issues of global developmental trends emphasize "a prime importance in sustainable development that is to meet the needs of the present without compromising the ability of future generations to

meet their own needs" (Zimmermann, 2003). Thus, sustainable personality, and, consequently, e-business application user, is "a person who sees relationships and inter-relationships between nature, society and the economy" (Rohweder, 2007). In other words, this is a person who is able to develop the system of external and internal perspectives, and in turn the system of external and internal perspectives becomes a main condition for the sustainable e-business application user to develop.

For instance, the concern of the European Union, namely, to become "the most competitive and dynamic knowledge-based economy in the world capable of sustainable economic growth with more and better jobs and greater social cohesion" (European Commission, 2004), demonstrates the significance of developing the system of external and internal perspectives for the development of humans, institutions, society and mankind. Thus, the life necessity to develop the system of two perspectives, namely, external and internal, determines the research methodology of e-business applications in engineering education, as highlighted in Figure 1.

However, in real life sustainable e-business application user is often realized from one of the perspectives: from the internal perspective accentuating cognition (Vossen, 2009) and from the external perspective accentuating social interaction and finding a balance between the external and internal perspectives (Surikova, 2007).

The methodological foundation of the present research on use of e-business applications of Web 2.0 in engineering education is formed by the System-Constructivist Theory based on Parson's system theory where any activity is considered as a system, Luhmann's theory which emphasizes communication as a system, the theory of symbolic interactionalism and the theory of subjectivism. The System-Constructivist Theory introduced by (Luhman, 1988, pp. 1–14) and (Parson, 1976) emphasizes that human being's point of view depends on the subjective aspect: everyone has his/her own system of external and internal perspectives (Figure 1) that is a complex open system (Rudzinska, 2008) and experience plays the central role in a construction process (Maslo, 2007).

3 E-BUSINESS APPLICATIONS

The paradigm change, namely, the move towards collaborative business - from person to people and from systems to service (Jones, 2008), puts the emphasis on the use of e-business applications of Web 2.0 technologies. Typical e-business applications of Web 2.0

techniques and technologies include corporate blogs, wikis, feeds and podcasts (Vossen, 2009).

Blogs are seen by Vossen (Vossen, 2009) as a common way to stay in touch with customers, to inform about new products and to receive immediate feedback; they can also be used internally in order to discuss specific topics among the staff of an enterprise, in particular if people are geographically distributed. Blogs allow a moderated interaction between participants, be it customers or colleagues, and a simple and efficient distribution of announcements, experiences, opinions, reports, or evaluations. However, bloggers need to keep in mind that blogs are typically crawled by search engines, so that company internals written into a blog might reach the outside world and be presented as search results. Also, a blog is useless without regular updates, a reasonable number of readers, continuous moderation, and good content. It is also a good idea for a company to treat independent bloggers just like regular journalists, since they might have a considerable readership.

In order to stay up-to-date with a company blog, but also with other information an enterprise might publish, there are essentially two approaches: pull and push. The active or pull way is to read the information at my own liberty and pace; in the passive or push approach, the information will be delivered to me automatically. Indeed, blog entries and other sources can be subscribed via feeds that are based on protocols such as RSS or Atom (Johnson, 2006), and they can be read using a feed reader such as Bloglines, Google Reader, Newsgator, or NewsAlloy (or simply in the browser). A podcast is determined by Vossen (Vossen, 2009) as a particular form of feed consisting of audio or video material. Wikis allow collaborative work on a common set of documents by many authors, and have been discovered as a new way of performing knowledge management in a learning organization. If staff members can be motivated to participate in the development of a wiki, this can be considered as a good example of making implicit knowledge explicit, thereby attacking the core problem of knowledge management.

A social network can also act as a means of connecting employees of distinct expertise across departments and company branches and help them build profiles in an easy way, and it can do so in a much cheaper and more flexible way than traditional knowledge management systems. Once a profile has been set up and published within the network, others can search for people with particular knowledge or expertise and connect to them. If the social network is to be run outside an enterprise, providers like Ning allow an easy setup of a self-regulated and self-managed

External Perspectives
as a Life Necessity

Internal Perspectives
as a Life Necessity

synthesis

• conscious
• learning
• foreign language
• systematic
• from abstract
   to concrete
• thought generalization

the system of
interacting
phenomena

• unconscious
• acquisition
• mother tongue
• non-systematic
• from concrete to
   abstract
• object generalization

synthesis
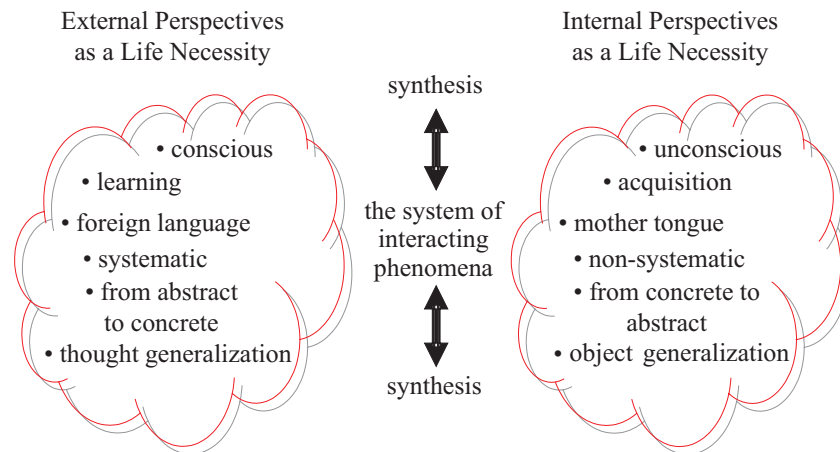
Figure 1: Developing the system of external and internal perspectives as a life necessity.

community. Vossen (Vossen, 2009) summarizes that a broad usage of Web 2.0 techniques and tools within an enterprise, paired with an increased exploitation of services offered over the Web and with leaving more room for the individual and its preferences, has led to the term "enterprise 2.0." In other words, enterprises as well as software vendors are exploiting it by integrating Web 2.0 features into their software, processes, and work environments. Hence, e-business applications of Web 2.0 technologies, namely, corporate blogs, wikis, feeds and podcasts, are seen as an integral part of engineering education.

4 E-BUSINESS APPLICATIONS IN ENGINEERING EDUCATION

The change in engineer entering the service area, namely, not working permanently at a large-scale enterprise but accepting project-related orders of large-scale enterprises by free engineers' office (Bassus and Wolfgramm, 2009) reveals the significance of e-business applications of Web 2.0 technologies to be integrated into the processes and environments of engineering education.

A proper integration of e-business applications into engineering education is provided by needs analysis. However, the emphasis of the System-Constructivist Theory on the subjective aspect of human being's point of view and experience that plays the central role in a construction process does not allow analyzing the student engineer needs objectively: human beings do not always realize their experience and their wants (Maslo, 2007).

In accordance with the research methodology, namely, developing the system of the external and in-

ternal perspectives, needs analysis is revealed to be of three levels, namely, individual, organizational and professional needs, in engineering education. Moreover, needs analysis includes four domains, namely, student's needs, wants, lacks and expectations, to analyze (Karapetjana, 2008). Thus, needs analysis has the potential to contribute decisively to the sustainable incorporation of e-business applications of Web 2.0 technologies or enterprise 2.0 into engineering education.

5 EMPIRICAL RESULTS

The target population of the present empirical study involves 22 participants of Fifth Baltic Summer School Technical Informatics and Information Technology at the Institute of Computer Science of the Tartu University, August 7-22, 2009, Tartu, Estonia and 40 students at the Department of Electrical Engineering and Computer Science of the Faculty of Engineering of Wismar University, University of Technology, Business and Design. All 22 participants of Fifth Baltic Summer School Technical Informatics and Information Technology have got Bachelor or Master Degree in different fields of Computer Sciences and working experience in different fields. The International Summer School offers special courses to support the internationalization of education and the cooperation among the universities of the Baltic Sea Region. The aims of the Baltic Summer Schools Technical Informatics and Information Technology are determined as preparation for international Master and Ph.D. programs in Germany, further specialization in computer science and information technology and learning in a simulated environment. The Summer School Technical Informatics and Information Tech-

nology contains a special module on Web 2.0 where e-business applications are an integral part.

Finally, 40 students at the Department of Electrical Engineering and Computer Science at the Faculty of Engineering of Wismar University, University of Technology, Business and Design were taken into consideration at the beginning of the seventh semester in the fourth year of their bachelor studies. The students have not got any or few work experience. The seventh semester of the Bachelor's program for Electrical Engineering and Computer Science at the Faculty of Engineering of Wismar University does not contain a special module on Web 2.0 where e-business applications are an integral part.

Analysis of the use of Web 2.0 is based on needs analysis as a basis for designing (Surikova, 2007) the following questionnaire:

- Question 1: Do you know the word Web 2.0?

- Question 2: Do you know the basic idea of Web 2.0?

- Question 3: Have you already used Web 2.0, namely, Facebook, Twitter, Wikipedia, etc?

- Question 4: Do you think Web 2.0 requires a lot of profound knowledge, namely, math, physics, etc?

- Question 5: Do you think Web 2.0 is useful for your individual needs?

- Question 6: Do you think Web 2.0 is useful for your organizational use?

- Question 7: Do you think Web 2.0 is useful for your professional use?

The evaluation scale of five levels for each question is given where "1" means "disagree" and low level of experience in use of e-business applications of Web 2.0 technologies and "5" points out "agree" and high level of use of Web 2.0. The participants' use of Web 2.0 was evaluated by the participants themselves on the first day of the Baltic Summer School, namely, August 7, 2009, and by the student engineers themselves at the beginning of the seventh semester in the fourth year of bachelor studies, namely, September 2009.

The analysis of the survey (Figure 2) reveals the following: the use of Web 2.0 by the Baltic Summer School (BaSoTi) participants is heterogeneous and the participants consider Web 2.0 to be most useful for their individual needs. The analysis of the survey (Figure 3) in September 2009 emphasizes that the university students' use of Web 2.0 is heterogeneous as well as the students don't know the possibilities offered by Web 2.0 properly. Then, the comparison of the survey results of the BaSoTi participants and the university students reveals the following:
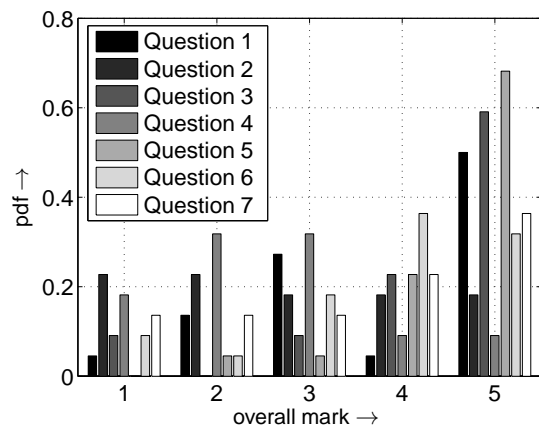
Figure 2: PDF (probability density function) of the BaSoTi participants' evaluation on August 7th, 2009.
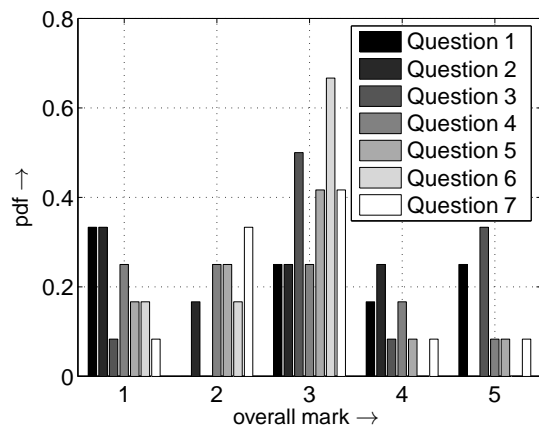
Figure 3: PDF (probability density function) of the university students' evaluation in September 2009.

- Question 1: The word Web 2.0 to a wider extent is known by the BaSoTi participants thereby developing the internal perspective that can be explained by a higher level of their education, namely, obtained Bachelor or Master Degree in different fields of Computer Sciences.

- Question 2: The basic idea of Web 2.0 to a higher degree is obtained by the BaSoTi participants as well thereby developing the internal perspective, that can be also revealed by a higher level of their education, namely, obtained Bachelor or Master Degree in different fields of Computer Sciences, and their participation in international projects, for example, Baltic Summer School Technical Informatics and Information Technology.

- Question 3: A wider experience in use of Web 2.0, namely, Facebook, Twitter, Wikipedia, is stressed by the BaSoTi participants that can be meant by the fact that the BaSoTi participants mostly come from the Baltic states where English is an interna-

tional means of communication whereas the university students use German that is already an international language thereby finding a balance between the external and internal perspectives.

- Question 4: Both groups, namely, the BaSoTi participants and the university students, consider that Web 2.0 does not require a lot of profound knowledge in math and physics, thereby developing the internal perspective that emphasizes the role of both groups as users and not as Web 2.0 designers.

- Question 5: The use of Web 2.0 for the individual needs is emphasized by the BaSoTi participants thereby developing the internal perspective that can be meant by the educational peculiarities, namely, the use of Web 2.0 in educational systems of the Baltic States.

- Question 6: The use of Web 2.0 for the organizational purposes is considered by the BaSoTi as well as university students, thereby finding the balance between the external and internal perspectives explained by the increased use of Web 2.0 in the educational system.

- Question 7: The use of Web 2.0 for the professional use is outlined by the BaSoTi participants thereby developing the system of the external and internal perspectives that can be also revealed by a higher level of their education, namely, obtained Bachelor or Master Degree in different fields of Computer Sciences, and their participation in international projects, for example, Baltic Summer School Technical Informatics and Information Technology.

Hence, the use of Web 2.0 and, consequently, e-business applications by the BaSoTi participants is provided by the knowledge the participants obtained in Bachelor or Master studies in different fields of Computer Sciences and by their working experience in different fields thereby putting the emphasis on developing the internal perspective while the use of Web 2.0 by the university students is regarded as finding the balance between the external and internal perspectives.

Due to the educational differences, namely, Web 2.0 module volume and content, between the Baltic Summer Schools Technical Informatics and Information Technology and the Bachelor's program for Electrical Engineering and Computer Science at the Faculty of Engineering of Wismar University in the seventh semester, Survey 2 of use of e-business applications of Web 2.0 technologies by only the Baltic Summer School participants was carried out on August 11, 2009.

Between Survey 1 and 2 of the participants'
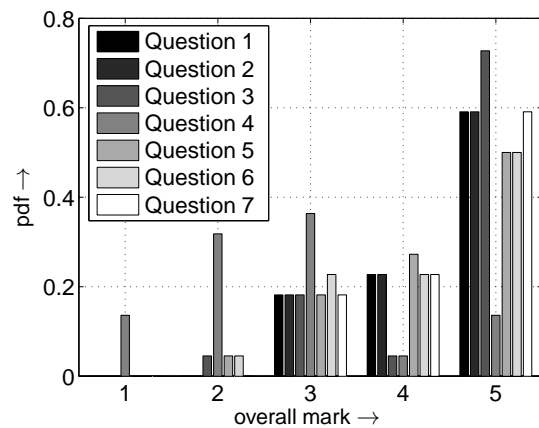
Figure 4: PDF (probability density function) of the second BaSoTi participants' evaluation on August 11th, 2009.

experience in use of e-business applications teaching/learning activity involved courses in Technical Informatics and Information Technology (German and English), preconference tutorials for introduction into advanced research topics, attendance of conference Advanced Topics in Telecommunication, tutorials and practical tasks, language training for talk and presentation (optional in English or German), leisure activities and social contacts, practical work at IT Company. Then, the analysis of the second survey (Figure 4) reveals that the participants' experience in use of e-business applications has become homogeneous and the participants have put the emphasis on use of Web 2.0 where e-business applications are an integral part for professional needs.

After having implemented a variety of methods and forms of teaching/learning activity (Zaščerinska, 2009) the result summary of two surveys of the participants' experience within the Baltic Summer School 2009 demonstrates the positive changes in comparison with Survey 1: the level of the participants' experience in terms of use of Web 2.0 has been enriched, the level of the participants' experience in terms of knowledge of basic idea of Web 2.0 has been improved, the level of the participants' experience in terms of use of Web 2.0 for individual needs decreased, thereby developing the system of the external and internal perspectives and the level of the participants' experience in terms of use of Web 2.0 for organizational and professional needs increased, thereby developing the system of the external and internal perspectives. Thus, the results' comparison of Survey 1 and Survey 2 of the participants' experience in use of e-business applications emphasizes the decrease of the participants' number who have obtained the low and critical level of experience and the increase of the participants' number who have achieved the average

and optimal level of experience revealed by the significance in difference between the levels of the participants' experience in use of e-business applications.

6 CONCLUSIONS

The results reveal that needs analysis of e-business applications in engineering education contributes to the incorporation of business-driven topics in the curriculum of engineering science, thereby developing the system of the external and internal perspectives.

The recommendation here is the role of teachers as mentors for participants' self-discovery and self-realization; to motivate participants, to stimulate their interests, to help them to develop their own structure and style, as well as to help them to evaluate their performance and be able to apply these findings to improve (Maslo, 2007) their further use of e-business applications in engineering education.

Further research on needs analysis of e-business applications in engineering education is considered to include needs analysis, the questionnaire development, carrying out empirical studies and statistical analysis.

REFERENCES

Bassus, O. and Wolfgramm, K. (2009). The Innovative Entrepreneurship Training Program at Hochschule Wismar - University of Technology, Business and Design. *First Asian Conference on e-Business and Telecommunications*, Berlin: Mensch & Buch, pp. 73-81.

European Commission (2004). Implementation of "Education and Training 2010" Work Programme. *Working Group B "Key Competences"*.

European Union (2008). *Lifelong Learning for Creativity and Innovation. A Background Paper*.

Johnson, D. (2006). *RSS and Atom in Action – Web 2.0 Building Blocks.*

Jones, S. (2008). The Move Towards Collaborative Business - From Person to People and From System to Service. *Capgemini Global Outsourcing*.

Karapetjana, I. (2008). *English for Specific Purposes Teaching Methodology*. PhD Thesis, University of Latvia.

Luhman, N. (1988). *Erkenntnis als Konstruktion*. Bern: Benteli.

Maslo, E. (2007) Transformative Learning Space for Life-Long Foreign Languages Learning. *International Nordic-Baltic Region Conference of FIPLV Innovations in Language Teaching and Learning in the Multicultural Context*, pp. 38–46.

Parson, T. (1976). *Theorie sozialer Systeme*. Opladen: Westdeutscher Verlag.

Rohweder, L. (2007). What kind of Sustainable Development do we talk about?, *Towards Sustainable Development in Higher Education - Reflections*. Helsinki University Press, Finland, pp. 22-27.

Rudzinska, I. (2008). The Quality of Aim Setting and Achieved Results in English for Specific Purposes-Study Course in Lecturers and Students' Opinion. *Proceedings of the ATEE Spring University Conference Teacher of the 21st Century: Quality Education for Quality Teaching*, pp. 366-374.

Surikova, S. (2007). Development of Criteria, Indicators and Level System for Evaluation of Enhancement of Primary School Students' Social Competence. *Society, Integration, Education*, pp. 383-393.

Vossen, G. (2009). Web 2.0: A Buzzword, a serious Development, just Fun, or What? *International Joint Conference on e-Business and Telecommunications (ICETE)*, 7.–10, pp. IS33–IS40.

Zaščerinska, J. (2009). Designing Teaching/Learning Activities to Promote e-Learning. In Ahrens, A. and Lange, C. (ed.), *First Asian Conference on e-Business and Telecommunications*, Berlin: Mensch & Buch, pp. 22–35.

Zimmermann, B. (2003). *Education for Sustainable Development–Baltic 21. An Agenda 21 for the Baltic Sea Region*. Danish Ministry of Education.

FROM DIGITAL ARCHIVES TO E-BUSINESS
A Case Study on Turning "Art" into "Business"

Rungtai Lin and Chih-Long Lin

Graduate School of Creative Industry Design, National Taiwan University of Arts, Taichung County, Taiwan
rtlin@ntua.edu.tw

Abstract: Along with Information Technology progress, E-business is becoming a key concept in the Internet and electronic commerce world. However, in today's intensely competitive business climate, innovative products become central to E-business development. Furthermore, changes in consumer perceptions regarding innovation are also important in E-business. Recently, creative industries are continually emerging in electronic commerce and have the potential to become a key trend in E-business. Understanding the E-business models for creative industries and helping designers to design "culture" into products are important research issues, and issues not yet well covered. Therefore, this paper proposes an ABCDE approach to illustrate how to transform "Archive" into "E-business". In order to turn "Archive" into "Business", we first need "Creativity" and "Design"; only then can we transform innovative products into "E-business." Results presented herein create an interface for looking at the way E-business crosses over cultures, and illustrate the interwoven experience of E-business and cultural creativity in the innovation design process and electronic commerce world.

1   INTRODUCTION

Along with the progress of Information Technology, E-business has become a common concept in the Internet and electronic commerce world. However, in today's intensely competitive business climate, innovative products become central in E-business development (Amit & Zott, 2001; Ben Lagha, Osterwalder & Pigneyr, 2001). To be successful, innovative products must have a clear and significant difference that is related to a market place need. Furthermore, changes in consumer perceptions regarding innovation are also important in E-business. In addition, "Culture" plays an important role in the design field, and "cross cultural design" will be a key design evaluation point in the future (Lin, 2007; Lin et al., 2007). Designing "culture" into modern product will be a design trend in the global market. E-business is considered to be one of the pivotal components in cultural and creative design industries, and this will have a significant impact on consumer perception of innovation.

In the global market - local design era, connections between culture and E-business have become increasingly close. For E-business, cultural value-adding creates the core of product value. It's the same for culture; E-business is the motivation for pushing the development of cultural and creative industries forward. Recently, creative industries have been continually emerging in electronic commerce and can become a key trend in E-business. Obviously, we need a better understanding of E-business in cultural and creative design industries, and not only for the global market but also for local design. While cross-cultural factors become important issues for product design in the global economy, the intersection of E-business and creative industries becomes a key issue making both local design and the global market worthy of further in-depth study (Dubosson, Osterwalder & Pigneyr, 2002; Osterwalder & Pigneyr, 2002)..

The importance of studying E-business has been shown repeatedly in several studies in various areas of the design field (Amit & Zott, 2001; Ben Lagha, Osterwalder & Pigneyr, 2001). Despite the recognized importance of E-business in cultural and creative design industries, industries lack a systematic approach to E-business. Understanding the E-business models for creative industries and turning "arts" into "business" for designers are important research issues, and until now these topics

have not been well covered. In order to transform "Archive" into "Business", we first need "Creativity" and "Design", only then can results be transformed into "E-business." (Ko et al., 2009). Therefore, this paper proposes an ABCDE approach for illustrating how to transform "Archive" into "E-business". The ABCDE approach integrates the difference between products and services of cultural and creative design industries into the E-business activities of current service development practice.

The ABCDE model provided illustrates how the National Taiwan University of Arts (NTUA) has established a link between E-business and cultural and creative industries through Our Museum, Our Studio and Our Factory respectively. Through the E-business approach, we have been able to merge design, culture, creativities and economy. The approach also further illustrates some other implications of the approach through the cultural perspective. Results presented herein create an interface for looking at the way E-business crosses over cultures, and illustrates the interwoven experience of E-business and cultural creativity in the innovation design process and electronic commerce world.

2 FROM OEM TO OBM

Taiwan's industrial design is developing along with its economic development. The design development could be represented as a smile face, proposed by the former ACER president Shi, from OEM (Original Equipment Manufacture), ODM (Original Design Manufacture), to OBM (Original Brand Manufacture) as shown in Figure 1 (Lin et al., 2007).
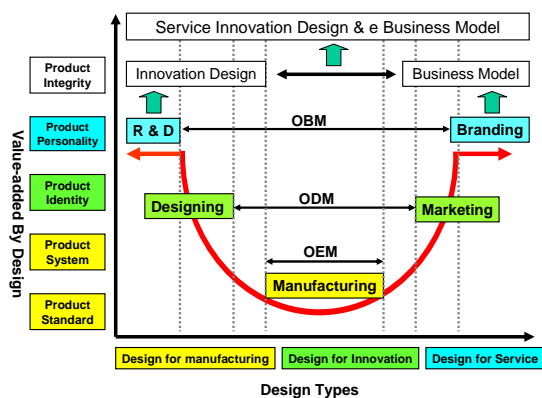
Figure 1: From OEM to OBM in e Business.

Before 1980, OEM vendors in Taiwan reduced costs to produce "cheap, high quality" products as a strategy to become successful in the global manufacturing industry. With the OEM style of having "cost" but without a concept of " price" in mind, or just by knowing "cost down" but not knowing " value up", these vendors created Taiwan's economic miracle by earning a low profit from manufacturing. Those dependent upon hard-working patterns from the OEM pattern became obstacles in developing their own design. These vendors were extremely busy producing products to meet manufacturing deadlines; there was no time to develop design capabilities, so that the environment could not nurture design talents (Lin, 2007; Lin et al., 2007).

After 1980, Taiwan enterprises began to develop ODM patterns to extend their advantages in OEM manufacturing. Taiwan's government addressed a series of measures to stimulate the nation's economic growth, including the "Production Automation Skill Guidance Plan", and the "Assisting Domestic Traditional Industrial Skill Plan". These plans were to guide vendors to make production improvements, to lower costs and to increase competition. Starting from 1989, the industry Bureau pushed the "Plan for Total Upgrading of Industrial Design Capability" over three consecutive five-year plans. The scheme established working models by experienced design scholars and students from universities for the purpose of working on design. The design students worked with the enterprises on specific projects to set up a working pattern of industrial design based on enterprises' real needs (Lin, 2007;Lin et al., 2007). Recently, product design in Taiwan has stepped into the OBM era. In addition, cultural and creative industries have already been incorporated into the "National Development Grand Plan", demonstrating the government's eagerness to transform Taiwan's economic development by "Branding Taiwan" using "Taiwan Design" based on Taiwanese culture (Lin, 2007; Lin et al., 2007).

There has been a recent shift from technological innovation to E-business based on discovering new opportunities in the marketplace. Companies are more focused on adapting new technologies and combining them in ways that create new experiences and value for customers. With the development of industrial trend, most companies gradually realized that the keys to product innovation are not only aspects of market and technology but also service innovation design (Baxter, 1995; Zhang et al., 2003). Ulrich and Pearson (1998) point out that service design has received increased attention in the academic and business communities over the past

decade. Both academics and practitioners emphasized that the role of service design in innovative product development relates not only to aesthetics, but also to aspects such as ergonomics, user-friendliness, efficient use of materials, and functional performance (Gemser & Leenders, 2001).

However, we now live in a small world with a large global market. While the market heads toward "globalization", design tends toward "localization." So we must "think globally" for the market, but "act locally" for design. While E-business is under tough competitive pressure from the developing global market, it seems that the local design should be focused on E-business in order to adapt innovation to product design (Gregoire & Schmitt, 2006).

3 CONCEPTUAL FRAMEWORK

After reviewing the development of Taiwan's industrial design, it is clear that E-business is the force pushing cultural and creative industries development forward. The main purpose of this paper is to study factors affecting the E-business model. These factors are discussed in order to understand E-business in cultural and creative design industries. Then, a conceptual framework is proposed for defining, classifying, assessing, and modeling the E-business model for the cultural and creative industries.
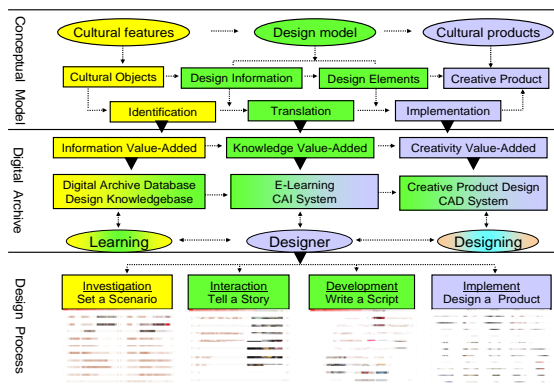
Figure 2: The conceptual framework .

The conceptual framework in Figure 2 consists of three main phases; *conceptual model*, *digital archive method*, and *design process*. The conceptual model focuses on how to extract cultural features from cultural objects and then transfer these features to the design model. The digital archive method consists of three phases; *identification* (extract cultural features from original cultural objects),

translation (transfer them to design information and design elements) and *implementation* in the final design of a cultural product (Lin, 2007; Lin et al., 2009).

3.1 Conceptual Model

The conceptual model is shown in the top of Figure 2 and includes three stages: *identifying cultural features*, *building the design model* and *designing cultural products*. To accomplish the goal, there are four steps including: *selecting cultural objects*, tr*ansforming design information*, *extracting design elements*, and *designing creative design products*. Then, implementation of the goal is broken into three phases: *identification*, *translation* and *implementation* which are described as follows (Lin, 2007; Lin et al., 2009):

Identification phase**:** the cultural features are identified from original cultural objects including the outer level of colors, texture, and pattern, the mid level of function, usability, and safety (Holzinger, 2005), and the inner level of emotion, cultural meaning, and stories (Heimgärtner, Holzinger & Adams, 2008). The designer uses the scientific method and other methods of inquiry and hence is able to obtain, evaluate, and utilize design information from the cultural objects.

Translation phase: the translation phase translates the design information to design knowledge within a chosen cultural object. The designer achieves some depth and experience of practice in these design features and at the same time is able to relate this design knowledge to design problems in modern society. This produces an appreciation for the interaction between culture, technology, and society.

Implementation phase: the implementation phase expresses the design knowledge associated with the cultural features, the meaning of culture, an aesthetic sensibility, and the flexibility to adapt to various designs. At this time, the designer gains knowledge of cultural objects and an understanding of the spectrum of culture and value related to the cultural object. The designer combines this knowledge with his strong sense of design to deal with design issues and to employ all of the cultural features in designing a cultural product.

3.2 Digital Archives Database

How to build a digital archive database is shown in the middle of Figure 2 and includes information value-added, knowledge value-added, and creativity value-added. The application of cultural features is a

powerful and meaningful approach to product design. Consumers nowadays require a design which is not only functional and ergonomic, but which also stimulates emotional pleasure. Lin (2007) took a cultural object called the Linnak as the example to build a digital archive database for learning culture through the internet and e-learning environment. The data collected after studying its appearance, usability, and cultural meaning is shown in Table 1. A design-related format was used to match the different items based on tribe, name of object, type, image, material, color, appearance, usability, pattern, form grammar, form structure, form style, inner content, and original resource. These items covered three levels of cultural characteristics and basic information such as imagery icon, tribe, and name. We propose that this information will serve as a reference for designers during the product design phase (2005; Hsu, 2004; Lin, Cheng & Sun, 2007).

Table 1: The format of the cultural features of Linnak.

Object	Linnak or twin cup
Type	Drinking container
Tribe	Paiwan, Rukai
Picture	
Material	Wood
Color	Natural wood color or painting with colors
Pattern	Figure, human-head, long-hooded pit viper pattern, Deer pattern
Principle of formation	1. Embossment on handles with a variety of patterns. 2. Total length from 43cm to 91cm, pitch from 29cm to 42cm, and cup capacity from 300 to 600c.c. 3. Single cup with a rectangular column shape and handle on both sides. 4.The Linnak contains two rectangular cups, a beam bridge in between and a handle on both sides.
Classification	Twin-cup, Single-cup and Tri-cup.
Operation	Two drinkers are required to hold the handles with left and right hand on each when drinking alcohol.
Using Scenario	1. Single cup is created only for the chief to drink liquid in the Paiwan tribe. Sometimes it was used to contain rice alcohol and to reward a hunting hero for demonstrating valor. 2. The twin -cup was created for use in wedding ceremonies where the bride and groom were required to drink alcohol together. 3. The Tri-cup was created for the groom and bride and chief (a witness), who stands between groom and bride to drink alcohol together which represents the greatest honor and wish for the couple.
Cultural content	1. The long- hooded pit viper or ancient figure pattern on the cup enhances the value of cup. 2. The twin-cup was mostly used in festival ceremonies to demonstrate a warm and harmonious spirit. 3. To drink with the twin cup represented the commitment in love between a male and female. 4. Drinking together represents eternal friendship.
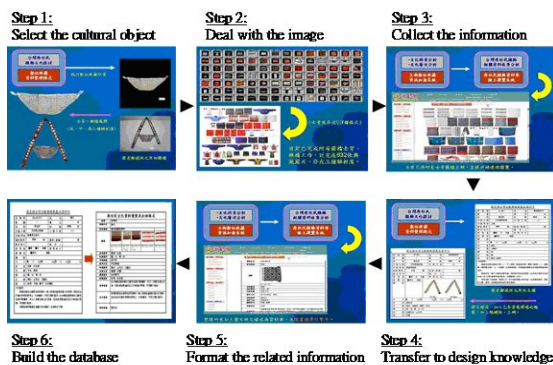
Figure 3: Process of building the database.

According to Table 1, a digital archive database was built to help to understand both the hard and soft contents of the cultural object. A process of building a digital archive database included six steps

(Figure 3): (1) select the cultural object, (2) deal with the image, (3) collect the information, (4) transfer the information to design knowledge, (5) format the related information, and (6) build the database. In addition, a friendly interface was provided to the designer for accessing the database easily as shown in Figure 4 and 5 (Hsu, 2004; Lin, Cheng & Sun, 2007).

Figure 4: Interface of the database.
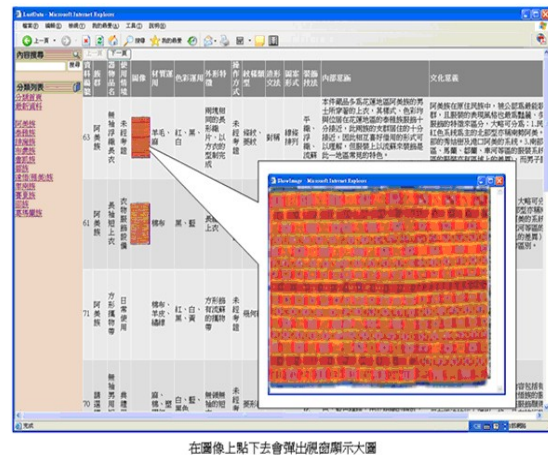
Figure 5: Interface for referring the pattern.

3.3 Design Process

Based on the cultural product design model, the cultural product is designed using scenario and story-telling approaches. In a practical design process, four steps are used to design a cultural product: *investigation* (set a scenario), *interaction* (tell a story), *development* (write a script), and *implementation* (design a product) as shown in the bottom of Figure 2 (Hekkert, 2003; Leong & Clark, 2003).

The four steps of the cultural product design process are described as follows:

Step 1 / Investigation / Set a scenario: The first

step is to find the key cultural features in the original cultural object and to set a scenario to fit the three levels: the outer 'tangible' level, the mid 'behavioral' level, and the inner 'intangible' level. Based on the cultural features, the scenario should consider the overall environment such as economic issues, social culture, and technology applications. This step seeks to analyze the cultural features in order to determine the key cultural features to represent the product.

Step 2 / Interaction / Tell a story: Based on the previous scenario, this step focuses on a user-based observation to explore the social cultural environment in order to define a product with cultural meaning and style derived from the original cultural object. Therefore, some interactions should be explored in this step, including interaction between culture and technology, dialogue between users and designers, and understanding the user's needs and cultural environment. According to the interaction, a user-centered approach is used to describe the user need and the features of the product by story-telling.

Step 3 / Development / Write a script: This step addresses concept development and design realization. The purpose of this step is to develop an idea sketch in text and pictograph form based on the developed scenario and story. During this step, the scenario and story might require modification in order to transform the cultural meaning into a logically sound cultural product. This step provides a means to confirm or clarify the reason why a consumer needs the product and rationale of how to design the product to fulfill the users' needs.

Step 4 / Implementation / Design a product: This step deals with previously identified cultural features and the context of the cultural products. At this point, all cultural features should be listed in a matrix table which will help designers check the cultural features in the design process. In addition, the designer needs to evaluate the product features, product meaning, and the appropriateness of the product. The designer may make changes to the prototype based on results from the evaluation, and implement the prototype and conduct further evaluations.

Based on the cultural product design model, Figure 6 shows how to transfer the original object -- 'Pottery-pot' from the Paiwan tribe into a design for a modern bag. Different cultures use textile containers designed for their own storage and transportation needs. Unlike bags or containers made from rigid materials such as clay or glass, textile containers offer flexibility of use by adapting to whatever item they are carrying. they also have the great advantages of being non-breakable and easy to

store. Figure 7 shows how to use the Taiwan aboriginal garments as the original cultural objects to design modern bags. In addition, Figure 8 demonstrates the cultural features extracted from Taiwan aboriginal garment culture and then transformed into modern bag design.

Figure 6: The modern bag designed from a pottery pot.

Figure 7: The process of designing modern bags from cultural objects.

Figure 8: Various bags from the Taiwan aboriginal garment culture.

4 E-BUSINESS MODEL

After explaining why business executives and academics should consider a rigorous approach to E-business models, we introduce a new E-business

Model for the cultural and creative industries. The new model is called "ABCDE Plan" which shows that to turn "Art" into "Business", we need "Creativity" and "Design", which allows the creative products to be transformed into "E-business" as shown in Figure 9.
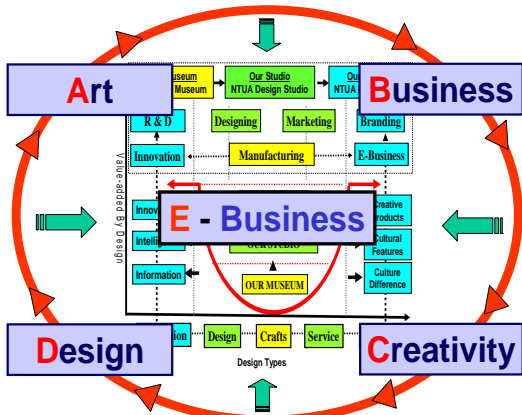
Figure 9: The concept of e-Business model.

To implement the ABCDE plan, National Taiwan University of Arts (NTUA) established an art museum, known as "Our Museum", in 2007 for the purpose of linking professional teaching with the museum's research, education, and display functions. At the same time the museum would present cultural and aesthetic ideas about art and artifacts to the public. Developing craftsmanship and creativity as well as competences related to the arts are of strategic importance to NTUA. Therefore, a design studio, known as "Our Studio", was subsequently set up in the College of Design in NTUA with the purpose of providing innovative products. NTUA is located in the Taipei metropolitan area, one of the most competitive regions in Taiwan. This area contains a significant concentration of craftsmanship and research establishments, linked by various formal and informal networks. Due to the challenging nature of the cultural and creative industries, NTUA is devoted to developing its regional and international networks by operating a cultural and creative industry park, known as "Our Factory." NTUA has established the link between "Art" and "Business" and has combined "Creativity" and "Design" through Our Museum, Our Studio and Our Factory respectively. It is a new approach that integrates design, culture, artistic craftsmanship, creativities and service innovation design in cultural and creative design industries (Roy & Riedel, 1976; Stevens, Burley & Divine, 1999).

With the increasing globalization of the economy, rapidly developing information technology, rapidly growing market competition, shortening life cycles of products and services, and increasing customer demands, companies and public sector actors will find it increasingly difficult to survive based on their past operating models. Therefore, based on the previous review of service design change, we propose a conceptual framework to innovation service design of cultural and creative design industries by using the smile paradigm as shown in Figure 10 (Ko et al., 2009; Lin et al., 2009).
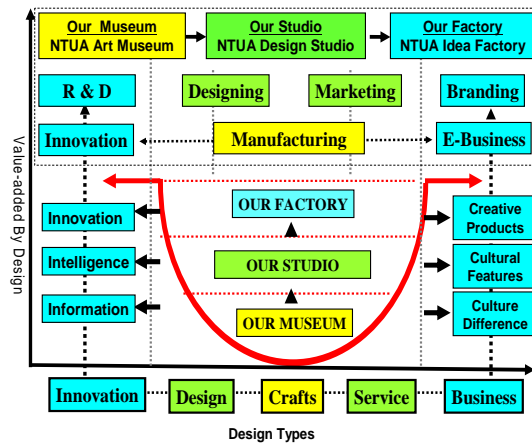
Figure 10: e-Business for creative industries.

According to the smile paradigm, craftsmanship is a part of Cultural creativity, and like the mouth in the smile face, it must still go up through innovation design and branding before it can become a "business". However, craftsmanship is not the entirety of culture, nor is creativity the whole of business; good craftsmanship at best earns outsourcing money, like an OEM vendor. The key to innovation design is to blend craftsmanship, creativity and service design, and "branding" is the key to any business (Ravasi & Lojacono, 2005).

In general, craftsmanship is the use of local materials to develop localized skills; localization is an important force behind the globalization of any international conglomerate, especially in the employment of cultural creativity. Crafted products produced in small volume seek to represent the spirit of "attention to details", and are a demand on the person, a representation of the person, an expression by the person, and a story from the person. Craftsmanship plumbs the depth of skills, while creativity seeks the height of impression, and branding asks for the width of acceptance. Only through culture and creativity, by allowing craftsmanship and creativity to facilitate branding,

can one makes one's way in this field (Yair et al., 1999, 2001;Veryzer, 1998; Voss & Zomerdijk, 2007).

The goal of the cultural and creative park is to combine artistic craftsmanship and economy with service design, and ultimately establish NTUA as a distinctive trademark associated with the park. To accomplish this goal, NTUA aims to combine artistic craftsmanship from "Our Museum" with cultural creativities from "Our Studio" in order to result in aesthetics in business for "Our Factory". Creativity and business are the elements for reaching an aesthetic economy. It is the concept of "Think Globally - Act Locally" to process the "Digital Archive" of Our Museum through the cultural creativities of Our Studio, producing cultural products in Our Factory in order to establish a local industry making aesthetic and economical products (Ko et al., 2009; Lin et al., 2009).

The current development of the Cultural Creative Park at NTUA is based on creative knowledge of crafts elements and materials from Our Museum and Our Studio. This cultural information is then transferred into the creative industry. In the near future, we will further implement this distinctive mode of cultural creative production to promote the concept of "Savoring Culture" which has the potential to become a "Taiwan industry concept". We are encouraging more and more creative products which contain colourful Taiwanese culture and styles. By supporting the development of cultural creative industry of NTUA, we can enjoy the fruitful success of an aesthetic culture in the creative industry (Lin et al., 2009).

5 CONCLUSIONS

With increasing global competition, E-business in service innovation design is not merely desirable for a company; it is a necessity. The importance of studying E-business as part of service innovation design has been shown repeatedly. However, there is no systematic approach that covers E-business in cultural and creative design industries. Therefore, a new approach was proposed by applying E-business in service innovation design in the domain of cultural and creative design industries. The E-business in service innovation design model is presented herein to provide designers with a valuable reference for designing "service" into a successful cross-cultural product. The purpose of this paper is to fulfill the aesthetic experience by connecting design and culture. This is turn will

synthesize technology, humanity, cultural creativities. Finally, we will achieve the aim of promoting service design amongst the general public.

For future studies, we need a better understanding of the acculturation process not only for the E-business in service design, but also for innovative product design. While cultural features become important issues in the interactive experiences of users, the acculturation process between human and culture becomes a key issue in cultural product design and worthy of further in-depth study. However, the effectiveness of using E-business in cultural and creative industries can be further enhanced. This can be done by incorporating more information of best practice in service industries into E-business in cultural and creative design industries.

ACKNOWLEDGEMENTS

REFERENCES

Amit, R. and Zott, C., 2001. Value Creation in E-business. *Strategic Management Journal*, Vol. 22, No. 6/7, Special Issue: Strategic Entrepreneurship: Entrepreneurial Strategies for Wealth Creation (Jun. - Jul., 2001), 493-520.

Ben Lagha, S., Osterwalder, A., Pigneur, Y., 2001. Modeling E-business with eBML*, 5e Conférence International de Management des Réseaux d'Entreprises*, Mahdia, October 25-26.

Berkley, B.J., 1996. Designing services with function analysis. *Hospitality Research Journal*, Vol. 20, No.1, 73-100.

Baxter M. 1995. *Product design: a practical guide to systematic methods of new product development.* Chapman & Hall, London, UK.

Dubosson, M, Osterwalder, A. & Pigneur, Y., 2002. eBusiness Model Design, Classification and Measurements, Preprint on an article accepted for publication in Thunderbird *International Business Review*, January 2002, vol. 44, no. 1: 5-23

Grégoire, B. & Schmitt, M., 2006. Business service network design: from business model to an integrated multi-partner business transaction. *Proceedings of the 8th IEEE International Conference on E-Commerce Technology and the 3rd IEEE, International Conference on Enterprise Computing, E-Commerce, and E-Services (CEC/EEE'06)*

Gemser G., & Leenders M., 2001. How integrating industrial design in the product development process impacts on company performance. *The Journal of Product Innovation Management 18*, 28-38.

Hekkert, P., Snelders, D., & van W ieringen, P. C. W., 2003. Most advanced, yet acceptable: Typicality and novelty as joint predictors of aesthetic preference in industrial design. *British Journal of Psychology*, *94(1)*, 111-124.

Holzinger, A. (2005). Usability Engineering for Software Developers. *Communications of the ACM,*48,1, 71-74.

Heimgärtner, R., Holzinger, A., Adams, R. (2008). From Cultural to Individual Adaptive End-User Interfaces: Helping People with Special Needs. In: Lecture Notes in *Computer Science* (LNCS 5105), (pp. 82–89). Berlin, Heidelberg, New York: Springer.

Hsu, C.H., 2004. An application and case studies of Taiwanese aboriginal material civilization confer to cultural product design. Unpublished master's thesis, Chang Gung University, Taoyuan, Taiwan.

Ko, Y.Y., Lin, P.H. & Lin, R., 2009. A Study of Service Innovation Design in Cultural and Creative Industry. *HCI (14) 2009*: 376-385.

Leong, D. & Clark, H.. 2003. Culture -based knowledge towards new design thinking and practice - A dialogue. *Design Issues*, 19(3), 48-58.

Lin, R., & Kreifeldt, J. G., 2001. Ergonomics in Wearable Computer. *International Journal of Industrial Ergonomics 27, Special Issue : Ergonomics in Product Design,* 259-269.

Lin, R., 2007. Transforming Taiwan Aboriginal Cultural Features Into Modern Product Design－A Case Study of Cross Cultural Product Design Model. *International Journal of Design*, 1(2), 45-53.

Lin, R., Cheng, R., & Sun, M. X., 2007. Digital Archive Database for cultural product design. HCI International 2007, 22-27 July, Beijing, P.R. China. paper- ID:825, *Proceedings Volume 10, LNCS_4559, ISBN: 978-3-540-73286-0.*

Lin, R., Sun, M.X., Chang, Y.P., Chan,Y.C., Hsieh,Y.C., & Huang,Y.C., 2007. Designing "Culture" into Modern Product --A Case study of Cultural Product Design. *HCI International 2007*, 22-27 July, Beijing, P.R. China. Paper-ID:2467, *Proceedings Volume 10, LNCS_4559, ISBN: 978-3-540-73286-0.*

Lin, R., Lin, P.B., Shiao, W.S. & Lin, S.H., 2009. Cultural Aspect of Interaction Design beyond Human-Computer Interaction. *HCI (14) 2009*: 49-58.

Martin, C.R. & Horne, D.A., 1993. Services innovation: Successful versus unsuccessful firms. *International Journal of Service Industry Management*, Vol. 4, No. 1, 49-65

Osterwalder, A., & Pigneur, Y., 2002. An E-business Model Ontology for Modeling E-business. 15th Bled Electronic Commerce Conference e-Reality: Constructing the e-Economy, Bled, Slovenia, June 17 - 19, 2002.

Ravasi, D., & Lojacono, G., 2005. Managing design and designers for strategic renewal. *Long range planning*, 38, 51-77.

Roy R., & Riedel J., 1997. Design and innovation in successful product competition. *Technovation 17,* 537-548.

Stevens, G., Burley, J., & Divine, R., 1999. Creativity + business disciplines = higher profits faster from new product development. *Journal of Product Innovation Management* 16, 455-468.

Ulrich, K. T., & Pearson, S., 1998. Assessing the importance of design through product archaeology. *Management Science 44*, 352-369.

Veryzer, R. W., 1998. Discontinuous innovation and the new product development process. *Journal of Product Innovation Management 15*, 304-321.

Voss, C. & Zomerdijk, L., 2007. "Innovation in Experiential Services – An Empirical View". In: DTI (ed). *Innovation in Services.* London: DTI. pp.97-134.

Wu, T.Y., Hsu, C.H., and Lin, R., 2004. The study of Taiwan aboriginal culture on product design, In Redmond, J., Durling, D. & De Bono, A. (Eds.), *Proceedings of Design Research Society International Conference.* Paper #238, DRS Futureground, Monash University, Australia.

Yair, K., Press, M. & Tomes A., 2001. Crafting competitive advantage: crafts knowledge as a strategic resource. *Design Studies, 22*, 377-394.

Yair, K., Tomes, A. & Press, M., 1999. Design through marking- crafts knowledge as facilitator to collaborative new product development, *Design Studies, 20*(6), 495-515.

Zhang, J., Tan, K.C., & Chai, K.H., 2003. Systematic Innovation in Service Design Through TRIZ. In the Proceedings of the EurOMA-POMS 2003 Annual Conference, Cernobbio, Lake Como, Italy, June 16th - 18th, 2003, Vol 1, pp. 1013-1022.

INTELLIGENT MOBILE SAFETY SYSTEM TO EDUCATIONAL ORGANIZATION

Li-Shan Chen

Department of Information Management, Fortune Institute of Technology
No. 1-10, Nongchang Rd., Daliao Township, Kaohsiung County 83160, Kaohsiung, Taiwan (R.O.C.)
sham@center.fotech.edu.tw

Abstract   This study aims to develop safety system, and let the system become intelligent. We adopt the swarm intelligence and active Ultra-High Frequency RFID for safety system, and develop friendly human-computer-interface software for users use the personal digital assistants. We program the system and software with Extensible Markup Language (XML) and C sharp language. If the users begin to search, the kernel safety system automatically communicates with other RFID readers by agents, and the agents can search the closer camera for users. This study's result has successfully implemented to one educational organization, and it would be helpful for the paterfamiliases to hold all situations about their children at the educational organization. That will be great help in the grip of whole after-school remedial education, teaching and learning situation. We select 100 paterfamiliases to test this system and software. It is revealed that 93% of the paterfamilias are satisfied with the system (Strongly agree：25/100; Agree：68/100; Disagree：5/100; Strongly disagree：2/100). The software searching correctness is 95% (Strongly agree：30/100; Agree：65/100; Disagree：3/100; Strongly disagree：2/100).

1   INTRODUCTION

Technological developments in content-based analysis of digital video information are undergoing much progress, with ideas for fully automatic systems now being proposed and demonstrated (Hyowon, Alan, Noel, & Barry, 2006). Effective agent teamwork requires information exchange to be conducted in a proactive, selective, and intelligent way (Fan, Wang, Sun, & Yen, 2006). Self-managing systems (i.e. those that self-configure, self-protect, self-heal and self-optimize) are the solution to tackle the high complexity inherent to these networks (Barco, Lázaro, Díez, & Wille, 2008) Digital representations are widely used for audiovisual content, enabling the creation of large online repositories of video, allowing access such as video on demand (Justin & Timothy, 2006). Digital artifacts created via transformational technologies often embody implicit knowledge that must be correctly interpreted to successfully act upon the artifacts (Leonardi & Bailey, 2008) With continued

advances in communication network technology and sensing technology, there is astounding growth in the amount of data produced and made available through cyberspace (Chen & Liu, 2006). Felfernig et al., (2009) focus on the first aspect and present an approach which supports knowledge engineers in the identification of faults in user interface descriptions. Chen (2009) adopts the Windows Media Player along the RTP/RTSP protocol in order to embed the mobile information system into the users' machines (personal digital assistants or smart phones), and provides a solution (including hardware solutions) to promote campus safety management. He also combines the swarm intelligence and Web Services to transform a conventional library system into an intelligent library system having high integrity, usability, correctness, and reliability software for readers (Chen, 2008, 2010). L. S. Chen, and S. L. Chen (2007) built the intelligent system and developed a knowledge base of the computer-parts. Jannach, Leopold, Timmerer, and Hellwagner (2006) present a novel, fully knowledge-based approach for

building such multimedia adaptation services, addressing the above mentioned issues of openness, extensibility, and concordance with existing and upcoming standards.

This study adopts the swarm intelligence and active Ultra-High Frequency RFID for safety system, and develop friendly human computer interface software for users use the personal digital assistants (PDAs). We program the system and software with Extensible Markup Language (XML) and C sharp language. If the users begin to search, the kernel safety system automatically communicates with other RFID readers by agents, and the agents can search the closer camera for users.

2 RELATED WORK

2.1 For Mobile Communication

Malek and Frank (2006) have focused on determining a near-optimal collision-free path because of its importance in robot motion planning, intelligent transportation systems, and any autonomous mobile navigation system. A spanning tree is based on the autoconfiguration of mobile ad hoc networks and a novel approach for efficient distributed address autoconfiguration (Li, Cai, & Xu, 2007). Pavlou, Huigang, and Yajiong (2007) build upon the principal–agent perspective to propose a set of four uncertainty mitigating factors—trust, Web site informativeness, product diagnosticity, and social presence. A neural network is trained to learn the correlations and relationships that exist in a dataset (Kaikhah and Doddament, 2006). Gao and Zhang (2008) have proposed an effective technique to determine the number and distribution of equilibria and a new supervised linear feature extraction technique for multiclass classification problems particularly suited to the nearest-neighbor classifier technique (Masip and Vitria, 2008). Wang and Chen (2008) present a new method for evaluating students' answer scripts using vague values, where the evaluating marks awarded to the questions in the students' answer scripts are represented by vague values. Payne (2008) examines the Web service paradigm from an open multiagent system perspective and contrasts the formally grounded knowledge-centric view of agents with a pragmatic declarative bottom-up approach adopted by Web services. The

location-based spatial queries having certain unique characteristics can be revealed, which traditional spatial query processing systems employed in centralized databases do not address (Ku, Zimmermann, & Wang, 2008). Lee and Wang (2009) present an ontology-based computational intelligent multi-agent system for Capability Maturity Model Integration (CMMI) assessment. Medium access control protocols have quality-of-service support—topology-independent link activation transmission scheduling—for mobile code-division multiple-access ad hoc networks (Su, Su, & Li, 2008). The context-aware query processing system enhances the semantic content of Web queries using two complementary knowledge sources: lexicons and ontologies (Storey, Jones, Sugumaran, & Purao, 2008). Yap, Tan, and Pang (2008) propose the Explaining BN Inferences (EBI) procedure for explaining how variables interact to reach conclusions.

2.2 For RFID Systems

Broekmeulen, and Donselaar (2009) suggest a replenishment policy for perishable products which takes into account the age of inventories and which requires only very simple calculations. Zhou (2009) takes a different perspective by modeling item-level information visibility in general. Delgado, Ros, and Vila (2009) present a system that is able to process the information provided by a Tagged World to identify user's behavior and to produce alarms in dangerous situations. Abad et al., (2009) present important advantages regarding conventional traceability tools and currently used temperature data loggers such as more memory, reusability, no human participation, no tag visibility needed for reading, possibility of reading many tags at the same time and more resistance to humidity and environmental conditions. Lee, and Chan (2009) propose a genetic algorithm to determine such locations in order to maximize the coverage of customers. Also, the use of RFID is suggested to count the quantities of collected items in collection points and send the signal to the central return center. Angeles (2009) looks at the perceived ability of components of IT infrastructure integration and supply chain process integration to predict specific radio frequency identification (RFID) system deployment outcomes exploration, exploitation, operational efficiency, and market knowledge creation.

2.3 For Swarm Intelligence

Tabu search and ant colony perform better for large-sized problems, whereas simulated annealing is optimal for small-sized problems and it is therefore essential that a maintenance scheduling optimizer can incorporate the options of shortening the maintenance duration and/or deferring maintenance tasks in the search for practical maintenance schedules. Allahverdi and Al-Anzi (2008) addressed a two-stage assembly flow-shop scheduling problem with a weighted sum of makespan and mean completion time criteria, known as bicriteria. The learners and lecturers agree that style-based ant colony systems can provide useful supplementary learning paths (2008). Ant colony intelligence (ACI) is proposed to be combined with local agent coordination in order to make autonomous agents adapt to changing circumstances, thereby yielding efficient global performance. This indicates that the ACO algorithm is an optional compromise strategy between preferable phase unwrapping precision and time-consuming computations.

3 METHODOLOGY

This study adopts the active Ultra-High Frequency RFID and swarm intelligence for safety system, and develop friendly human computer interface software for users use the personal computers or notebooks. The system is developed in the environment of: Microsoft Windows Server 2008, Internet Information Services 7.0 (IIS 7.0), Microsoft Structured Query Language (MS SQL) Server 2008, and Visual Studio 2008 (VS 2008). The programming languages are Extensible Markup Language (XML) and C#.

3.1 RFID System Framework

Figure 1 shows the framework of RFID system.

3.2 Searching Design

The searching path of this study is from (Agents Generator) to (Reader)$_i$ and the distance of (Agents Generator) to (Reader)$_i$ is the closest, as shown in Figure 2. The searching path begins at (Agents Generator), and it has two choice. One is (Reader)$_1$, and the other is (Reader)$_2$. Because the

Figure 1: Framework of RFID system.

Figure 2: Searching Path.

distance of (Agents Generator) to (Reader)$_1$ is shorter than (Agents Generator) to (Reader)$_2$, the optimum searching path is (Agents Generator) to (Reader)$_1$. And so forth, the optimum searching path is "(Agents Generator) →(Reader)$_1$ →(Reader)$_3$ →(Reader)$_2$→…". Figure 2 shows the searching process. This study amends the ant algorithm of Birattari, Pellegrini, and Dorigo (2007). It can let the safety system become intelligent and mobility. The design of the "Agents Generator" is very important, and it is the kernel technology in this study. The developing process is described as below. (Note: The agents are seemed as the ants.)

3.3 Meaning of the Symbols and Nouns

(a) n: The numbers of RFID readers
(b) m= $\Sigma_{k=1 \text{to} n}$
   b_i (t): The total agents
   b_n (t): The numbers of agents in the (Agents Generator)
(c) d_{ij}: The distant of (Agents Generator) to (Reader)$_i$; This study considers that it is

(d) $\tau_{ij}(t)$: The intensity of pheromone upper edge

$$\tau_{ij}(t) = \rho\tau_{ij}(t) + \Delta\tau_{ij} \qquad (1)$$

This study uses (Eq. 1) to update the pheromone.

ρ：The parameters of pheromone evaporation

$$\Delta\tau_{ij} = \quad \Sigma_{k=1} \text{ to m } \Delta\tau_{ij}{}^{k} \qquad (2)$$

(e) $\Delta\tau_{ij}{}^{k}$: The kth agent remains pheromone going through the edge (i, j). It is defined as equation 3.

Q：The influential parameter of the pheromone

L_K ： The total length of the route, and the kth agent goes all over the (Readers)

$\Delta\tau_{ij}{}^{k} = Q / L_K$, The Kth agent goes through edge (i, j) between time point t and $(t + t_i)$

$$\Delta\tau_{ij}{}^{k} = 0, \text{ Otherwise} \qquad (3)$$

(f) R：The cycles counter agent goes through all of the readers, and the R_{max} is the upper limit of R

(g) $Tabu_k(I)$: The record of the kth has gone through the re, anaders the "I "is to mak a visit to "Ith"reader. It can prevent the agent from going back to cities already visited.

(h) μ_{ij}: The inverse of the distance of (Agents Generator) to (Reader)$_i$

$$\mu_{ij} = 1 / d_{ij} \qquad (4)$$

(i) $P_{ij}{}^{k}(t)$: The probability that kth agent goes from (Agents Generator) to (Reader)$_i$

Set $\Phi = k \in (n - Tabu_k (I))$
$P_{ij}{}^{k}(t) = 【\tau_{ij}(t)】^{\alpha} 【\mu_{ij}】^{\beta} / \Sigma_{\Phi} 【\tau_{ik}(t)】^{\alpha} 【\mu_{ik}】^{\beta}$ (5)
if $j \in (n - Tabu_k (t))$
Otherwise $P_{ij}{}^{k}(t) = 0$ (6)

The α and β are the important controlled parameters of pheromone information and μ_{ij}.

3.4 Designing Steps

The designing steps are described below.
Step 1: Set t=0、R= 0 （"t" is the time counter, and "R" is the cycles counter）For all edge (i, j), Set $\tau_{ij}(t)$ = Constant, $\Delta\tau_{ij}(t) = 0$. To put m agents into n readers
Step 2: Set I=1 ("I" is Tabu list index). For k = 1 to m (The record of the kth agent is listed in $Tabu_k(I)$ at agents generator.)
Step 3: Set I= I +1. For k = 1 to m (Using equation

5 to decide (Reader)$_i$ and moving the kth agent to (Reader)$_i$ recorded in $Tabu_k(I)$.)
Step 4: For k = 1 to m do .To move the kth agent from $Tabu_k(n)$ to $Tabu_k(1)$ and calculate the total length of all paths recorded, and update the shortest path. To calculate each edge (i, j). For k = 1 to m do. $\Delta\tau_{ij} = \Delta\tau_{ij} + \Delta\tau_{ij}{}^{k}$
Step 5: By $\tau_{ij}(t + t_1) = \rho\tau_{ij}(t) + \Delta\tau_{ij}$, Calculates $\tau_{ij}(t + t_1)$ for each edge (i, j). Set $t = t + t_1$, R = R + 1 for each edge (i, j). Set $\Delta\tau_{ij} = 0$ for each edge (i, j)
Step 6: If (R < R_{max}) and (No entering in stop situation) Then clear the entire Tabu list . Go To Step 2. Else print the shortest path and stop.

4 RESULTS AND DISCUSSION

4.1 Results

The safety system has been successfully developed, as shown in Figure 3-8. Figure 3 is the "login frame", and Figure 4 is "Welcome frame". There are eight areas in this system; (1) The left side of the gate, (2) The right side of the gate, (3) The eastern side of the house, (4) House Back, (5) Classroom 1, (6) Classroom 2, (7) Office, (8) Leisure area, see as Figure 5. Figure 6, Figure 7, and Figure 8 are searching results. For example, the user "missyang" logins to the system. The system will tell her that she has three children in the educational organization. One is in the classroom 1,

Figure 3: Login Frame.
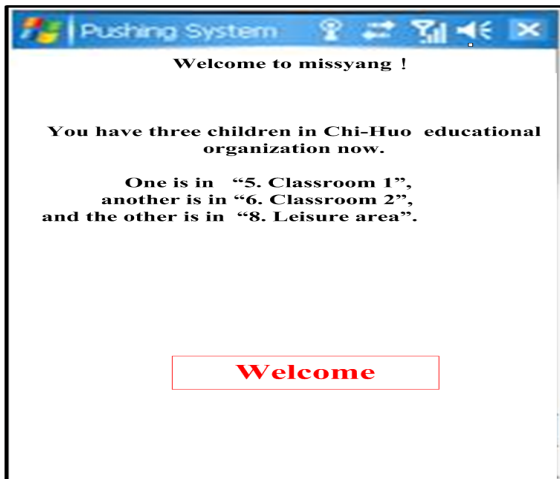
Figure 4: Welcome Frame.

Figure 5: Eight area.

Figure 6: Searching Results.

Figure 7: Searching Results.

another is in the classroom 2, and the other is in the leisure area. The missyang can push the "Welcome" key, and enter the choosing frame "Figure 5". She can choose and push the ⑤, ⑥, or ⑧ to watch her children's situations.
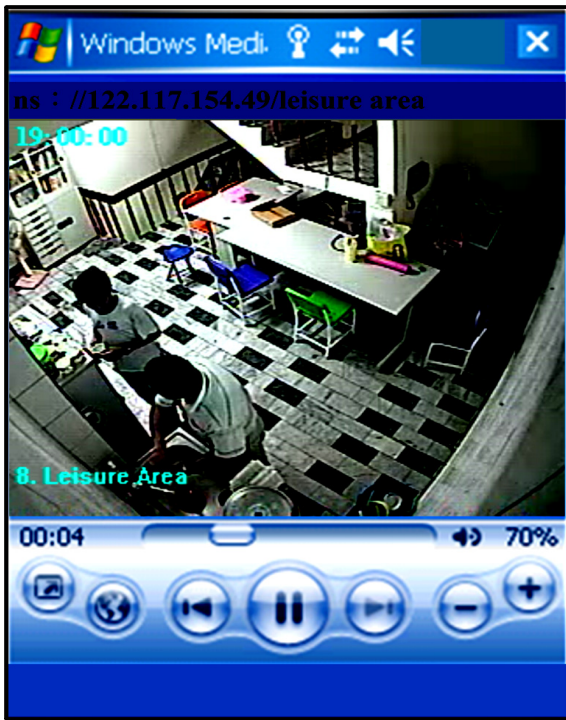
Figure 8: Searching Results.

4.2 Discussion

In this study, we select 100 paterfamiliases to test this system and software. It is revealed that 74% of the paterfamilias are satisfied with the system (Strongly agree：25/100; Agree：68/100; Disagree：5/100; Strongly disagree：2/100), see as Figure 9 and Figure 10.

The software searching correctness is 95% (Strongly agree ： 30/100; Agree ： 65/100; Disagree：3/100; Strongly disagree：2/100), see as Figure 11 and Figure 12.
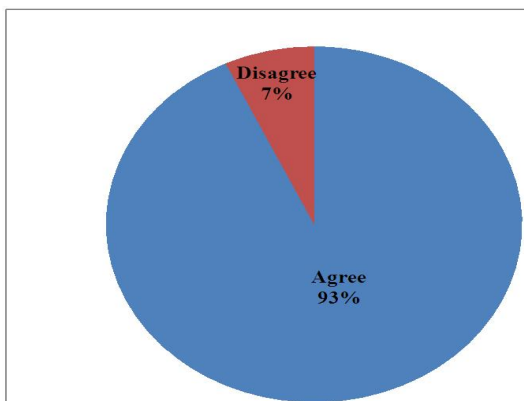
Figure 9: Satisfaction and Dissatisfaction.
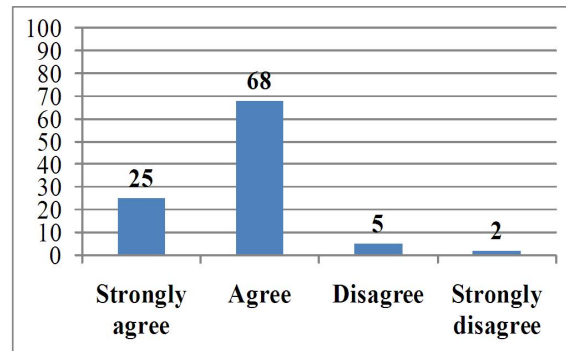
Figure 10: Detail of satisfaction and dissatisfaction.
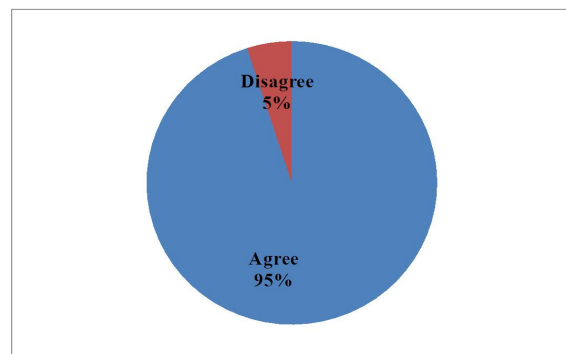
Figure 11: Software searching correctness.
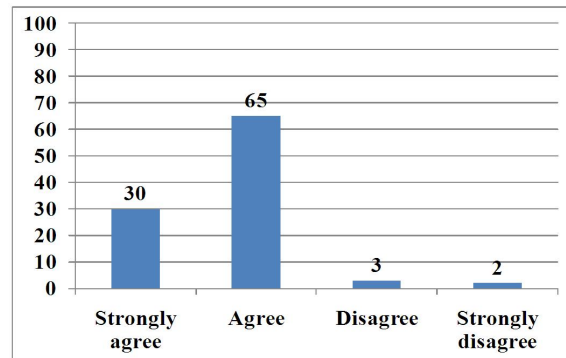
Figure 12: Detail of software searching correctness.

5 CONCLUSIONS

This study aims to developed safety system, and let the system become intelligent. This study used artificial intelligence and active Ultra-High Frequency RFID directly to guide paterfamilias monitoring their children's in-time images. Thus, it could save the paterfamilias' time on operating the

instrument. Even someone who has not the professional knowledge about information technology could use them skillfully. This study also develops friendly human computer interface software for users use the personal digital assistants. The size of the software is 22 kilobits; therefore, the software is not a liability for the users' tools. This study has successfully implemented to one educational organization, and it would be helpful for the paterfamiliases to hold all situations about their children at the educational organization. That will be great help in the grip of whole after-school remedial education, teaching and learning situation.

ACKNOWLEDGEMENTS

REFERENCES

Abad, E., Palacio, F., Nuin, M., Zárate, A. G., Juarros, A., Gómez, J. M., & Marco, S. (2009). RFID smart tag for traceability and cold chain monitoring of foods: Demonstration in an intercontinental fresh fish logistic chain. Journal of Food Engineering, 93 (4), 394-399.

Allahverdi, A., & Al-Anzi, F. S. (2008). The two-stage assembly flowshop scheduling problem with bicriteria of makespan and mean completion time. International Journal of Advanced Manufacturing Technology, 37 (1-2), 166-177.

Angeles, R. (2009). Anticipated IT infrastructure and supply chain integration capabilities for RFID and their associated deployment outcomes. International Journal of Information Management, 29 (3), 219-231.

Barco, R., Lázaro, P., Díez, L., & Wille, V. (2008). Continuous versus discrete model in autodiagnosis systems for wireless networks. IEEE Transactions on Mobile Computing, 7 (6), 673-681.

Birattari, M., Pellegrini, P., & Dorigo, M. (2007). On the invariance of ant colony optimization. IEEE Transactions on Evolutionary Computation, 11 (6), 732-742.

Broekmeulen, R. A. C. M & Donselaar, K. H. V. (2009). A heuristic to manage perishable inventory with batch ordering, positive lead-times, and time-varying demand. Computers & Operations Research, 36 (11) , 3013-3018.

Chen, K., & Liu, L. (2006). iVIBRATE: Interactive visualization-based framework for clustering large datasets. ACM Transactions on Information Systems (TOIS), 24 (2), 245-294.

Chen, L. S. (2008). Design and implementation of intelligent library system. Library Collections, Acquisitions, and Technical Services, 32 (3-4), 127-141.

Chen, L. S. (2009). Design and implementation of intelligent mobile information system for campus safety management. International Journal of Organizational Innovation, 2, 320-340.

Chen, L. S. (2010). Applying swarm intelligence to a library system. Library Collections, Acquisitions, and Technical Services, 34 (1), 1-10.

Chen, L. S., & Chen, S. L. (2007). Collaborative Design and Manufacture on Intelligent System. Journal of the Chinese Society of Mechanical Engineers, 28(2), 233-242.

Delgado, M., Ros, M., & Vila, M. A. (2009). Correct behavior identification system in a Tagged World. Expert Systems with Applications, 36 (6), 9899-9906.

Fan, X., Wang, R., Sun, S., & Yen, J. (2006). Context-Centric Needs Anticipation Using Information Needs Graphs. Applied Intelligence, 24(1), 75–89.

Felfernig, A., Friedrich, G, Isak, K., Shchekotykhin, K., Teppan, E., & Jannach, D, (2009), Automated debugging of recommender user interface descriptions. Applied Intelligence, 31(1), 1–14.

Gao, B., & Zhang, W. (2008). Equilibria and their bifurcations in a recurrent neural network involving iterates of a transcendental function. IEEE Transactions on Neural Networks, 19 (5), 782-794.

Hyowon, L., Alan, F. S., Noel, E. O'., & Barry, S. (2006). User evaluation of Físchlár-News: An automatic broadcast news delivery system. ACM Transactions on Information Systems (TOIS), 24 (2), 145-189.

Jannach, D., Leopold, K., Timmerer, C., & Hellwagner, C. H. (2006). A knowledge-based framework for multimedia adaptation. Applied Intelligence, 24 (2), 109–125.

Justin, Z., & Timothy, C. H. (2006). Detection of video sequences using compact signatures. ACM Transactions on Information Systems (TOIS), 24 (1), 1-50.

Kaikhah, K., & Doddament, S. (2006). Discovering Trends in Large Datasets Using Neural Networks. Applied Intelligence, 24 (1), 51-60.

Ku, W. S., Zimmermann, R., & Wang, H. (2008). Location-based spatial query processing with data sharing in wireless broadcast environments. IEEE Transactions on Mobile Computing, 7 (6), 778-791.

Lee, C. K. M., & Chan, T. M. (2009). Development of RFID-based Reverse Logistics System. Expert Systems with Applications, 36 (5), 9299-9307.

Lee, C. S., & Wang, M. H. (2009). Ontology-based computational intelligent multi-agent and its application to CMMI assessment. Applied Intelligence, 30(3), 203–219.

Leonardi, P. M., & Bailey, D. E. (2008). Transformational Technologies and The Creation of

New Work Practices ：Making Implicit Knowledge Explicit in Task-Based Offshoring. MIS Quarterly, 32 (2), 411-436.

Li, L., Cai, Y., & Xu, X. (2007). Spanning-tree based autoconfiguration for mobile ad hoc networks. Wireless Personal Communications, 43 (4), 1465-1477.

Malek, M. R., & Frank, A. U. (2006). A mobile computing approach for navigation purposes web and wireless geographical information systems. Lecture Notes in Computer Science 4295, 123-134.

Masip, D., & Vitria, J. (2008). Shared feature extraction for nearest neighbor face recognition. IEEE Transactions on Neural Networks, 19 (4) , 586-595.

Pavlou, P. A., Huigang, L., & Yajiong, X. (2007). Understanding and mitigating uncertainty in online exchange relationships: A principal—agent perspective. MIS Quarterly, 31 (1), 105-136.

Payne, T. R. (2008). Web services from an agent perspective. IEEE Intelligent Systems, 23 (2), 12-14.

Wang, H. Y., & Chen, S. M. (2008). Evaluating students' answerscripts using vague value. Applied Intelligence, 28 (2), 183–193.

Storey, V. C., Jones, A. B., Sugumaran, V., & Purao. S. (2008). A methodology for context-aware query processing on the World Wide Web. Information Systems Research, 19 (1), 3-25.

Su, Y. S., Su, S. L., & Li, J. S. (2008). Topology-independent link activation scheduling schemes for mobile CDMA ad hoc networks. IEEE Transactions on Mobile Computing, 7 (5), 599-616.

Yap, G. E., Tan, A. H., & Pang, H. H. (2008). Explaining inferences in Bayesian networks. Applied Intelligence, 29(3), 263–278.

Zhou, W. (2009). RFID and item-level information visibility. European Journal of Operational Research , 198 (1), 252-258

APPLYING MODEL-VIEW-CONTROLLER (MVC) IN DESIGN AND DEVELOPMENT OF INFORMATION SYSTEMS
An Example of Smart Assistive Script Breakdown in an e-Business Application

Andreas Holzinger, Karl Heinz Struggl

Institute of Information Systems and Computer Media (IICM), TU Graz, Graz, Austria
a.holzinger@tugraz.at, kstruggl@student.tugraz.at

Matjaž Debevc

Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia
matjaz.debevc@uni-mb.si

Abstract: Information systems are supporting professionals in all areas of e-Business. In this paper we concentrate on our experiences in the design and development of information systems for the use in film production processes. Professionals working in this area are neither computer experts, nor interested in spending much time for information systems. Consequently, to provide a useful, useable and enjoyable application the system must be extremely suited to the requirements and demands of those professionals. One of the most important tasks at the beginning of a film production is to break down the movie script into its elements and aspects, and create a solid estimate of production costs based on the resulting breakdown data. Several film production software applications provide interfaces to support this task. However, most attempts suffer from numerous usability deficiencies. As a result, many film producers still use script printouts and textmarkers to highlight script elements, and transfer the data manually into their film management software. This paper presents a novel approach for unobtrusive and efficient script breakdown using a new way of breaking down text into its relevant elements. We demonstrate how the implementation of this interface benefits from employing the Model-View-Controller (MVC) as underlying software design paradigm in terms of both software development confidence and user satisfaction.

1 INTRODUCTION

The process behind the production of a motion picture, be it a documentary, an action movie or an animation film, is typically segmented into the four phases of development, pre-production, production (or principal photography) and post-production. Analogous to other fields where project management is applied, the first phases of the film production management (FPM) process not only define and affect all efforts and results involved in the whole production, but also essentially decide whether the project will eventually be green-lighted (Clevé, 2005).

The producer's main responsibility in these first defining phases is to formulate the project (in terms of finding a screenplay script, acquiring licenses and possibly receiving commitment by important staff)

and to obtain funding for the actual realization of the project. For the latter task, it is vital that the producer create robust cost estimations. Contrary to later phases, where actual budgets are available (Singleton, 1996), these first cost estimations need to be based on cost defining elements in the script, which are identified and quantified by conducting script breakdown (Singleton, 1991).

This paper will first present the results of an observation on script breakdown interfaces of major film production management software suites and motivate the need for an unobtrusive and efficient script breakdown interface. Following this, the model-view-controller (MVC) pattern and its relevance to usability by software design will be discussed.

The proposal of auto-advancing shortkey (AAS) tagging will then follow and lead to a discussion of

how MVC was adopted in the system design phase and supported by the development environment. An observation of lessons learned and benefits of the provided solution will conclude this work.

2 SCRIPT BREAKDOWN IN THE FILM PRODUCTION PROCESS

Script breakdown is the act of analyzing the script of a film production (e.g. a movie screenplay), identifying its defining, relevant elements, and grouping them into categories (Clevé, 2005). For example, all speaking roles in a script are identified and assigned the "Actors" category.

The resulting breakdown information is, as noted, used to create first cost estimations. This is usually done by consulting so-called labor rate books and price lists provided by guilds and manufacturers, and applying them to the elements of the breakdown. For example, if in a given scene there are five extra cast members noted (e.g. to represent passengers waiting for a train), a labor rate book would be consulted that lists typical costs for such silent performers according to any possible special requirements, such as age or shooting conditions. Thus, it is essential to build such cost estimations on correct script breakdown data.

Furthermore, so-called breakdown sheets (Singleton, 1991) are created for every scene, providing a tabular summary of all contained elements, along with other production notes about the time of day the scene takes place, and similar. These breakdown sheets are of the utmost importance for the scheduling (time-planning) of the project. Also in later phases, they are considered the main reference document by all personnel involved in the preparation and shooting of scenes.

Historically, script breakdown has been done by printing out the script and highlighting elements with textmarkers that are color-coded according to their respective category (Singleton, 1991). Obvious drawbacks of this solution are the time required to conduct the breakdown and to transfer the data into breakdown sheets, scheduling plans, etc., as well as its error-proneness.

In recent years, film production software applications started adopting the idea of script breakdown and providing interactive interfaces for it. However, due to deficiencies in the solutions provided by these applications, many producers still resort to script printouts and textmarkers, and transferring and updating data manually.

Analysis of the script breakdown interfaces of five major film production and management software suites (CeltX, Final Draft and Final Draft

Tagger, Cinergy MPPS, Movie Magic Scheduling, Movie Magic Screenwriter) provided the following results:

1. Most often supported was fully manual tagging, i.e. tagging by manually selecting or entering text and category, either with or without the use of a separate tagging dialog window (all suites).

2. Tagging selected text into categories by context menu or category-buttons was supported by three suites.

3. One suite allowed to assign categories to the selected text by shortcut keys (e.g. "a" for actors).

4. Two suites provided a breakdown sheet-preview that could be manipulated interactively.

Two conclusions were drawn from these findings. First, most observed suites did not rely on one particular tagging mechanism, but provided up to three different interfaces. While this fact in itself does not disqualify any of the solutions, it shows that software design must be flexible and able to support different workflows with various involved interface elements independently of the underlying data model.

In the light of this observation, it must be stated that some issues concerning this requirement were discovered. This surfaced e.g. in the use of separate (unnecessary) dialog windows in three suites, or the fact that solutions using two or more windows only allowed the use of one window at a time, while the others were disabled. The system architecture presented later in this paper will employ MVC and point out how these problems can effectively be avoided by software design.

Second, all provided interfaces required the user to carefully select (or worse, manually input) text in order to tag an element with a category. This will obviously become difficult with typically dozens of pages long film scripts and therefore presents a likely reason for user frustration. The tagging interface proposed in this work will provide smart pre-selection of text in order to mitigate this problem.

3 METHODS AND MATERIALS

3.1 Auto-advancing Shortcut Key Tagging (AAS)

The idea of AAS is based on the findings presented in section 2. It combines the two main principles of auto-advancing using smart text pre-selection

heuristics, and the use of shortcut keys for assigning categories.

The pre-selection heuristics are responsible for finding the range of text that is most likely to be selected by the user, and, as the name implies, pre-selecting it, effectively mitigating the problem stated before. The heuristics themselves should be defined and implemented at one single point and therefore usable for various kinds of workflows. These include e.g. triggering after user has tagged an element (the program automatically selects the next likely element, i.e. auto-advancing) as well as normal navigation.

The rules implemented by the heuristics have to be carefully designed and defined to suit the actual use case. In the context of script breakdown, one important rule is that the user will never select single characters or parts of words for tagging. Therefore, the smallest granularity for selection is a word, i.e. upon auto-advancing (e.g. after tagging), the heuristics will select the next word after the current selection.

In order to provide more meaningful pre-selections, the heuristics should consider previously tagged elements, which may consist of multiple words, and include language-independent recognition of word boundaries.

The latter is a non-trivial task due to e.g. context-dependent meanings of punctuation (as shown in Figure 1) and language-specific ways of defining word boundaries.

The quick ("brown") fox can't jump 32.3 feet, right?

Thequickbrownfoxcan'tjump32.3feetright

Figure 1: Segmentation of an English sentence by word boundaries. Original sentence at the top, extracted version at the bottom. (Image extracted from (Unicode, 2009)).

Additionally, more complex language-specific processing can be included for better results, e.g. using specialized stopword-filters or word dictionaries. Assume, for example, the following text fragment as part of a movie script:

```
the Tin Man carries an axe
```

Assume also that the word `the` is currently selected by the user. Upon advancing, the pre-selection heuristics will select the word `Tin` for the user. Now, assume the same text fragment, but with the `Tin Man` already tagged, e.g. as speaking role:

```
the Tin Man carries an axe
```

In this case, the pre-selection heuristics would recognize the words Tin Man as compound token and select both for the user. Next, assume the user advanced once more:

```
the Tin Man carries an axe
```

In this last case, stopword-filtering can be applied to skip over the word an as it is reasonably safe to assume it will never represent an important element in an English script.

Shortcut key tagging of elements into categories on the other hand is a more user-centric design task. It requires a simple and flexible interface for managing categories and assigning them shortcut keys (and possibly formatting information, such as colors and underline- or highlighting-styles).

Both the pre-selection and the shortcut key tagging mechanisms inherently require the software to implement workflows that interact with several interface elements and controllers to reflect the user's actions in the software and in its data model.

3.2 Development Environment and Tools

The prototype described in the following sections was implemented on and for Apple Macintosh machines running Mac OS X 10.5 or above.

Design and implementation were done using the Xcode development environment, which is freely provided by Apple for any Mac OS X owner or member of the Apple Developer's Connection (ADC).

The underlying data model was built using Xcode's data model builder. All object persistence is managed by its Core Data module. Interface design was done using Interface Builder.

Classes and frameworks of the Mac OS X-native API Cocoa were used and partly subclassed and extended to implement the actual functionality and user interaction. It is to note that Xcode and the higher-level Cocoa API frameworks themselves employ MVC and support and advocate its use as design pattern.

3.3 Prototype Software

The prototype was developed with MVC as main underlying design pattern. MVC is an architectural software design pattern and engineering concept described by Trygve Reenskaug (Reenskaug & Skaar, 1989) (see Figure 2).
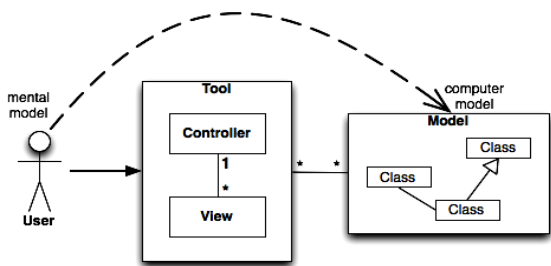
Figure 2: The general idea of the Model-View-Controller (MVC) paradigm, altered according to Reenskaug (Reenskaug, 2008).

3.3.1 Model

Analysis of the typical structure and the characteristics of movie scripts (as e.g. provided in (Clevé, 2005)) yielded the data model depicted in Figure3.
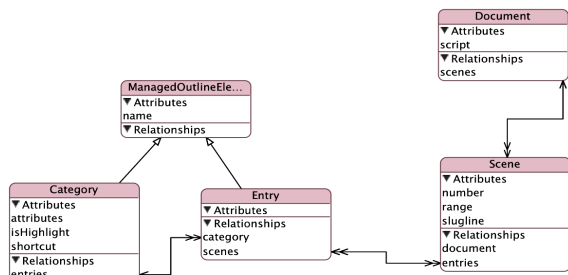
Figure 3: The prototype's data model as abstracted from real world entities.

Category and *Entry* share a common super-class that provides persistence and usage as outline items in the interface, as well as attributes and methods both subclasses need to. Category additionally administers a relationship to its elements (i.e. Entry objects representing text added to this category) and on how they are formatted in the script view. Entry on the other hand knows which Category it belongs to, and in which Scenes its specific text is tagged.

Scene objects are used to optionally limit the scope of tagging to scene boundaries. Also, they store references to all entries within them for the purpose of creating structure breakdown sheets and similar. Scenes are recognized and pre-processed automatically when a script is loaded into the breakdown prototype. The *Document* class is used for storing project information in order to enable saving and loading of projects including their scripts and all tagged elements.

It is to note that the entities defined in the data model above are represented in code by their respective model classes, e.g. *Category* is implemented in a *Category* class in the *Model*

directory and implements properties and methods pertaining only to itself. Inter-class workflows or interface functionality is implemented by controllers.

3.3.2 View

As noted, interface design was done using Xcode's Interface Builder. All controls and interface elements shown on the screen are implemented by so-called view-classes which are responsible for drawing themselves in order to create the user interface, as well as holding references to any controllers they may incorporate. E.g., a text view not only draws text (i.e. view code), but also provides text manipulation, which is handled by the views' text controller.

Where required, view and controller classes were subclassed to incorporate new or override existing behavior. E.g., the text view class was extended to handle shortcut keys and trigger pre-selection correctly after the according button presses.

Figure 4 gives an impression of the prototype's script view and category management interfaces. The main window supports tagging by AAS and alternatively by the use of a tagging context menu. The category management window features a scene selection menu, an outline view of all categories and their respective elements, and controls for manipulation and saving/loading category configurations.
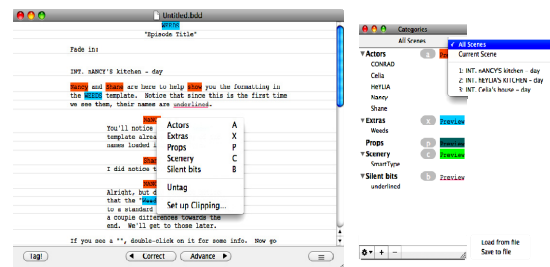
Figure 4: Breakdown prototype's main window and category management window including context menus and scene selection.

3.3.3 Controller

Two controllers in the breakdown prototype are responsible for implementing workflows and providing other general functionality across interface elements and involving several data model classes. Figure 5 shows how responsibilities between the two controllers are organized, including workflows and management of data model objects.
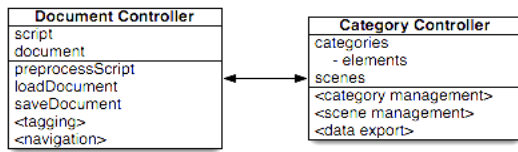
Figure 5: Responsibilities of the prototype's controllers.

The *Document Controller* is the core unit of the breakdown prototype as it implements saving and loading of projects and manages *Document* instances. It also encapsulates pre-processing of the script (for scene scanning) and tagging of elements using AAS (including pre-selection), the context menu or the *Tag*-button.

The *Category Controller* provides an interface to the *Document Controller* for tagging and handles user initiated category management itself. It also features a scene selection menu which is automatically generated from the scene data gathered by the *Document Controller*'s script pre-processing and can be used to bind tagging to a scene's scope, as well as to jump to the beginning of the scene in the script view.

Both the scene menu and the actual tagging process therefore present valid examples of how the system architecture effortlessly allows workflows to span over different controllers, interfaces and data model entities.

4 LESSONS LEARNED

It was noted before that employing MVC offers benefits for software systems dealing with a range of data model objects that the user can manipulate using an interactive interface. In order to give the user an impression of direct manipulation, as often discussed in software usability engineering, the prototype was implemented in a way that no user interaction or interface elements (e.g. tagging dialogs) are required that are not inherently part of the respective workflow (contrary to several of the observed film production software suites).

For example, tagging an element using a shortcut key first invokes an event on the text view, informing it about the user's keystroke. The subclassed view knows to forwards the event to its delegate (i.e. the controller class handling such user input, in this case the document controller) for handling.

The controller then invokes the AAS algorithm which itself uses a native library for word boundaries scanning. For every tagged element, the category controller is called with the required information to create an element object.

In short, this example shows that no artificial interface class or any inclusion of workflow logic in model classes is required. The document controller interprets and handles input as notified by the script view and invokes the category controller for creating the appropriate model class objects. The same applies e.g. for the script pre-processing step where the document controller scans for scenes and invokes the category controller to create scene elements.

Without the use of MVC, it is likely the system design would need to incorporate artificial pseudo-controllers to handle such workflows. In the system architecture presented herein, there are only two controllers responsible for a cohesive group of views and interactions using a clearly defined interface to implement complex workflows.

Another important aspect is reusability. For example, the algorithm for tagging a particular text is also used for un-tagging the text (i.e. removing from the category and un-formatting all occurrences in the text) and for re-formatting all of a category's occurrences if the category formatting itself has been changed. This is supported by the system design automatically since no additional user interaction (e.g. tagging dialog) is required. For the user, this means that un-tagging an element can be done by either selecting the element in the script (using AAS this is done semi-automatically) and hitting the backspace key, or by deleting the element from the categories outline view.

5 CONCLUSIONS

Work on the prototype has shown, as discussed, that MVC strongly advocates a system design that efficiently reflects workflows, business logic and therefore user interaction and its expected results in a way that is both intuitive for the user and plausible for the system engineer.

It can be said that MVC helped in minimizing the compromises between system usability and ease of development. In other words, most of the features and usability considerations of the prototype could have been taken into account and realized with other software development paradigms, but likely with the introduction of tradeoffs towards either system design or usability (as, again, was seen in several of the observed film production software suites).

Recommendable future work includes research of other, possibly similar software development paradigms for interactive applications and evaluation of how well they would suite the prototype at hand. For example, the Document-Context-Interaction (DCI) architecture recently proposed by Trygve

Reenskaug suggests a possibility for improvement over MVC by introducing roles to integrate generically defined workflows and business logic into the software design model (Reenskaug, 2008), (Reenskaug & Coplien, 2009).

REFERENCES

Bass, L. & John, B. E. (2003) Linking usability to software architecture patterns through general scenarios. *Journal of Systems and Software,* 66, 3, 187-197.

Booch, G. (1994) *Object-Oriented Analysis and Design with Applications.* Redwood City (CA)*,* Benjamin/Cummings.

Clevé, B. (2005) *Film Production Management.* Burlington, USA, Oxford, UK*,* Focal Press.

Curry, E. & Grace, P. (2008) Flexible self-management using the model-view-controller pattern. *IEEE Software,* 25, 3, 84-90.

Dahl, O.-J. & Nygaard, K. (1966) SIMULA: an ALGOL-based simulation language. *Communications of the ACM,* 9, 9, 671-678.

Holzinger, A. (2005) Usability Engineering for Software Developers. *Communications of the ACM,* 48, 1, 71-74.

Juristo, N., Moreno, A. M. & Sanchez-Segura, M.-I. (2007) Analysing the impact of usability on software design. *Journal of Systems and Software,* 80, 9, 1506-1516.

Kay, A. C. (1993) The early history of Smalltalk. *The second ACM SIGPLAN conference on History of programming languages.* Cambridge, Massachusetts, United States, ACM.

Kristaly, D. M. & Moraru, S. A. (2006). *Java technologies for model-view-controller architecture*. 10th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM 2006), Brasov (Romania), Transilvania Univ Press-Brasov, 175-178.

Leff, A. & Rayfield, J. T. (2001). *Web-application development using the Model/View/Controller design pattern*. 5th IEEE International Enterprise Distributed Object Computing Conference, Seattle, Wa, Ieee Computer Soc, 118-127.

Lethbridge, T. C. (2000) What Knowledge Is Important to a Software Professional? *IEEE Computer,* 33, 5, 44-50.

McLaughlin, B., Pollice, G. & West, D. (2006) *Head First Object-Oriented Analysis and Design.* Sebastopol (CA)*,* O'Reilly.

Oestereich, B. (1999) *Developing Software with UML: Object-Oriented Analysis And Design In Practice.* Harlow (UK)*,* Addison Wesley.

Reenskaug, T. (2008), The Common Sense of Object Orientated Programming. Online available: http://heim.ifi.uio.no/~trygver/themes/babyide, last access: 2010-02-20

Reenskaug, T. & Coplien, J. O. (2009), The DCI Architecture: A New Vision of Object-Oriented Programming. Online available: http://www.artima.com/articles/dci_vision.html, last access: 2010-02-10

Reenskaug, T. & Skaar, A. L. (1989) An environment for literate Smalltalk programming. *Conference proceedings on Object-oriented programming systems, languages and applications.* New Orleans, Louisiana, United States, ACM.

Seffah, A. & Metzker, E. (2004) The obstacles and myths of usability and software engineering. *Communications of the ACM* 47, 12, 71-76.

Singleton, R. S. (1991) *Film Scheduling, or, How Long Will It Take To Shoot Your Movie?* New York*,* Lone Eagle Publishing.

Singleton, R. S. (1996) *Film Budgeting, or, How Much Will It Cost To Shoot Your Movie?* New York*,* Lone Eagle Publishing.

Unicode (2009) Unicode Standard Annex 29. *Unicode Text Segmentation.*

REVIEWING THE E-COLLABORATION MARKETPLACE
A Survey of Electronic Collaboration Systems

Bettina Schauer, Michael Zeiller and Doris Riedl

Information Technology and Information Studies, University of Applied Sciences Burgenland
Campus 1, Eisenstadt, Austria
{bettina.schauer,michael.zeiller,doris.riedl}@fh-burgenland.at

Abstract: Electronic collaboration systems that support and enable communication, coordination and collaboration between people in shared projects, processes and teams within organisations and for cross-organisational use have significantly changed under the influence of Web 2.0 technologies and social software. The electronic collaboration marketplace is made up of numerous systems that offer a large variety of features. A classification approach is presented that classifies electronic collaboration systems and thus structures the diverse collaboration marketplace. Collaboration systems are evaluated and compared using a set of evaluation criteria that allow for the assessment of all major collaboration tasks. Thus completeness of systems as well as the main focus of applicability of individual collaboration systems is determined.

1 INTRODUCTION

Web 2.0, social software and Enterprise 2.0 are major trends that shaped information technology throughout the last years. Collaborative software has also been massively influenced by those concepts and technologies. Tools previously denoted under the terms *Computer Supported Cooperative Wor*k (CSCW) or *groupware* gained new momentum. The market of CSCW and groupware changed and converted to a marketplace of various kinds of electronic collaboration systems.

The appearance of social software – briefly defined as "software that supports group interaction" (Shirky, 2002 in Allen, 2004) – was one of the major driving forces in the change of the collaboration marketplace. In 2006 McAfee coined the term Enterprise 2.0 in his trend-setting paper "Enterprise 2.0: The Dawn of Emergent Collaboration" pointing out how companies can benefit from Web 2.0 technologies to support their knowledge workers (McAfee, 2006a). McAfee (2006b) defines Enterprise 2.0 as "... the use of emergent social software platforms within companies, or between companies and their partners or customers". However, the most important application of social software in the enterprise is to support group interaction and group collaboration.

Due to this impact of social software new tools and new vendors appeared on the market of collaborative software. Renowned "groupware" vendors reorganized their portfolio and introduced new social software features in their products. A number of open source solutions appeared as well. This led to a marketplace of high complexity and diversity containing a lot of different kinds of electronic collaboration systems.

In this paper we present an approach to classify electronic collaboration systems and thus to structure the entire marketplace. Collaboration systems are evaluated and compared using a set of evaluation criteria. Criteria are chosen that cover the entire range of collaboration tasks and interaction processes (the 4Cs communication, cooperation, collaboration, connection; supplemented by cross-sectional features). We focus on complete solutions of collaboration systems that support multiple kinds of social interaction.

In section 2 we narrow down the group of electronic collaboration systems we want to discuss and present related market studies and classification schemes. Section 3 introduces our approach of evaluating E-Collaboration systems and specifies a set of evaluation criteria. In section 4 the evaluation approach is demonstrated by comparing the results on a number of collaboration systems. Section 5 concludes this paper.

2 ELECTRONIC COLLABORATION

Electronic Collaboration - short: E-Collaboration or eCollaboration – is operationally defined by Kock (2005) in a general way as "collaboration using electronic technologies among different individuals to accomplish a common task". According to Kock (2005) research on Electronic Collaboration should include research on *Computer-mediated Communication* as well as research on *Computer Supported Cooperative Work* (CSCW) (Wilson, 1991). However, we will focus on computer-supported Electronic Collaboration with the help of E-Collaboration systems.

2.1 e-Collaboration Systems

Riemer (2007) describes E-Collaboration systems as "software for supporting communication, coordination and cooperation between people processes in groups". Several synonyms are often used interchangeably for this category of information systems: groupware, CSCW systems, collaborative software, cooperation systems. Riemer's definition is based on the basic types of social interaction that can be found in CSCW systems and groupware: communication – coordination – cooperation (Teufel et al., 1995). In a similar way Cook (2008) uses four primary functions to classify social software: communication – cooperation – collaboration – connection.

Communication allows people to converse with others and exchange information with the help of synchronous (e.g., chat, conferencing tools) and asynchronous (e.g., email, weblog, microblogging) communication tools (Riemer, 2009; Cook, 2008). Communication can be differentiated by medium – chronology – group of people.

Coordination allows a temporal or issue-related matching and agreement on tasks and resources. Typical operations of coordination support team members in coordinating appointments, processes and tasks in projects, plus surveys and ad-hoc workflow management.

Collaboration is a working practice whereby individuals work together on a non-routine cognitive task to achieve a common purpose (Austin and Burton, 2004; AIIM, n.d.). Collaboration encourages people to work with each other on particular problems, with shared commitment and goals (Cook, 2008). Collaboration tools encompass for example wikis and whiteboards, application sharing and desktop sharing. Collaboration takes advantage of the services of communication and coordination.

Cook (2008) adds another primary function to social software and Enterprise 2.0 tools: connection. Networking technologies enable people to make connections with and between both content and other people. Social networking is the most prevailing technology for connection, but there are also a number of enabling technologies like people profiling and people search.

In terms of this market analysis E-Collaboration systems are defined as software for supporting and enabling communication, coordination and collaboration between people in shared projects, processes and teams within organisations and for cross-organisational use (following Riemer, 2009). Tasks in Electronic Collaboration can be assigned to one of the primary interaction processes (the 4Cs):

- Communication
- Coordination
- Collaboration
- Connection

Complete E-Collaboration systems have to support all four types of social interaction.

2.2 Classification of e-Collaboration Systems

There exist several scientific and commercial market studies on E-Collaboration systems that structure and organize available software packages into system classes and categories and set up descriptive criteria. Riemer (2007) provides an in-depth analysis of E-Collaboration systems using cluster analysis. Riemer's catalogue of classification criteria is made up of six categories: group processes (communication, coordination, collaboration), usage of system (continuous, situational), role for group (primary, secondary), types of communication (e.g., text/voice/video messages, email, voice/video call, text/voice/video conference), shared resources and features (e.g., forum, surveys, application sharing, group calendar, whiteboard), awareness (informal, group structural, social, workspace). Using a cluster analysis Riemer identifies five system classes: everyday systems, integrated systems, coordination systems, meeting systems and specialized tools.

Illik (2009) structures E-Collaboration systems into five categories that are arranged in layers:

- Live communication (top)
- Extended team communication
- Basic team communication
- Team repository

■ Knowledge management (bottom)

Essential features are located at the bottom. Higher layers increase effectiveness and efficiency in teams.

The Forrester Wave™: Collaboration Platforms, Q2 2009 evaluates 11 vendors against more than 60 criteria that are arranged in three major groups: current offering, strategy, market presence (Koplowitz, 2009). The biggest group is *current offering* that is made up of seven groups: collaboration platform, language support, architecture and administration, monitoring and reporting, security, cross-platform support, information workplace readiness.

In contrary to theses studies our evaluation approach focuses on complete E-Collaboration solutions only (i.e., similar to integrated systems according to Riemer, 2007) that have to cover collaboration, coordination, communication and connection features to a certain extent. To allow for an objective evaluation of E-Collaboration systems the evaluation criteria and their weighting is based on a standardized reference use case that represents typical collaboration tasks. For specific scenarios this approach can be easily adapted (especially using individual weights, see section 3.3) to meet the particular needs of the customer.

3 EVALUATION OF E-COLLABORATION SYSTEMS

The evaluation of E-Collaboration systems involved the following three steps.

1) **Selection of e-Collaboration Systems.** Potential candidates for evaluation were selected according to predefined criteria which each tool had to fulfil in order to be defined as a complete E-Collaboration system. The result of this step was a list of E-Collaboration systems for evaluation.

2) **Definition of Evaluation Criteria.** The features and functionalities offered by E-Collaboration Systems were assigned to the four categories of social interaction processes – the 4Cs communication, cooperation, collaboration, connection. This step resulted in a feature list grouped by the 4C categories.

3) **Assessment.** Each E-Collaboration system was evaluated according to the functionalities belonging to the four categories. The result of the evaluation process was an assessment of E-Collaboration systems comparing their strengths and weaknesses

according to the 4C categories complemented by supporting categories.

3.1 Selection of e-Collaboration Systems

In a first step before the actual evaluation, tool candidates for the assessment had to be selected. This step involved defining criteria a tool had to fulfil to be regarded as an E-Collaboration system.

Each tool had to provide at least some functionalities to support all four types of social interaction summarised as the 4Cs. Thus the support of information processing, communication as well as coordination were defined as "must have" criteria. Collaboration software had to offer information processing features as well as communication facilities (at least asynchronous communication services like email) as well as coordination features, e.g., a group calendar, to be regarded as a complete E-Collaboration system. Furthermore an E-Collaboration system should offer functionalities for information sharing as well as synchronous live communication.

The reason for this pre-selection was that the assessment should contain complete systems for E-Collaboration that cover a variety of features instead of specialised solutions like just a wiki or a weblog. The selection process resulted in a list of about 40 software packages that qualified for the assessment. Among these tool candidates were the products of common vendors like Microsoft Office SharePoint Server 2007, IBM Lotus, Oracle Beehive and Oracle Collaboration Suite as well as open source products like for example Zimbra Collaboration Suite, phpGroupWare, OpenGroupware, Novell Open Workgroup Suite or Alfresco Share. The review also included systems of the visionaries as defined by the Gartner survey (2009) like Jive, Telligent or Socialtext Collaboration Platform.

3.2 Evaluation Criteria According to the 4Cs of Social Interaction

Based on reviews from literature, analysis of case studies and related market studies of E-Collaboration systems a feature list of all typical functionalities offered by such systems was established. In order to assess the E-Collaboration systems according to their strengths and weaknesses for certain applications these functionalities were grouped into the four categories of social interaction depending on their primary support (see section 2.1):

■ Communication

- Coordination
- Collaboration
- Connection

The category *communication* (16 features in total) was split into the sub-categories synchronous communication (instant messaging, conferencing functionalities, telephone, etc.), asynchronous communication (email, blogs, comments, etc.) and social presence. The category *coordination* covers features (18 in total) for task management, project management, workflows, organisation of meetings and appointments. Features for shared content production such as wikis or whiteboards and for content administration like shared folders, shared documents, versioning or tagging were assigned to the category *collaboration* (25 features in total), see Table 2 (section 4). These features were complemented by supporting technologies for shared content production like social tagging, social bookmarking and social cataloguing and by administrative services. The category *connection* comprises functionalities (7 in total) such as e.g. social networks, people search or people profiling.

In addition to these 4C categories some more characteristics of E-Collaboration systems were considered to be important for the evaluation. These criteria included functions that do not belong to one of the 4C categories but support all of them. Such functionalities (16 in total) include e.g. newsfeeds, personalisation, alerts, configurable areas, mashups, search, filtering, rating, documentation, (online) help and the support of handheld deliveries. Administration of E-Collaboration systems was also an important aspect of the evaluation in addition to the 4C criteria. Some of the administrative features assessed (10 in total) were user management, scalability, configuration, integration into existing systems and backup or recovery.

3.3 Assessment of e-Collaboration Systems

For the assessment the features of the various categories were weighted according to their importance within the category. The weights were assessed for a standard team collaboration scenario. In case of a specific application the weights have to be adjusted according to the particular needs. Each E-Collaboration system was assessed according to whether it supported a feature or not. In case the feature was supported the weight of this feature was added to the score of the E-Collaboration system in the respective category, in case the feature was not supported it did not increase the score.

The assessment resulted in a score for each category with a maximum of 100% per category as the evaluation was not undertaken for a specific use-case and thus all 4Cs were supposed to be equally important. Thus it is possible to compare E-Collaboration systems according to their suitability for supporting the 4C categories. The total of 100% per category was split up into scores representing the relative importance of the subcategories. The scores of the subcategories were again split up into scores for each functionality – see Table 2.

In addition to the quantitative assessment qualitative data was collected for each E-Collaboration system. The data included facts about the system like name, vendor, version and which operating systems, browsers or databases are supported as well as a valuation of the general strengths and weaknesses of the E-Collaboration system.

4 RESULTS

In the following, due to space restrictions, the results of five examples out of the 40 E-Collaboration systems evaluated in November & December 2009 are presented, deliberately not including the market leaders Microsoft and IBM.

- **Alfresco Share** an open source tool for enterprise content management by Alfresco Software Ltd
- **Socialtext** an enterprise social software by Socialtext Incorporated
- **Zimbra Collaboration Suite** an open source collaboration application by Zimbra
- **Jive Social Business Software** an enterprise communication and collaboration platform by Jive Software.
- **PHProject** an open source groupware suite by Mayflower.

Table 1 shows the scores that each system achieved in the categories communication, coordination, collaboration and connection as well as cross sectional functionalities and administration out of 100% as the total score for each category.

Socialtext offers an intuitive user interface combined with a lot of functionality that is highly integrated into the features offered by this suite. It provides new technologies like microblogging or mashups. Thus the strengths of Socialtext lie in the categories communication and connection. The main focus is on the social aspect by transparently connecting people with the corresponding content.

Table 1: Evaluation of five E-Collaboration systems.

	Alfresco Share	Socialtext	Zimbra	Jive SBS	PHPro-ject
Communication	18%	74%	63%	55%	51%
Coordination	29%	25%	65%	30%	83%
Collaboration	70%	69%	59%	51%	61%
Connection	50%	86%	58%	78%	58%
Cross sectional functionalities	59%	95%	51%	76%	55%
Administration	66%	87%	67%	76%	72%

Zimbra offered most functionalities in supporting coordination activities and provides an easy to use interface with a familiar look and feel. Its strengths are extensive search options, tagging for all sorts of content and the synchronisation with handheld deliveries. Webservices can be integrated into the workspaces by so called Zimlets. Zimbra offers good support for collaboration in small teams but it lacks on overview of the whole organisation and all corresponding projects.

Jive SBS´s strength is the category connection. It offers many features for building employee communities using social networking concepts. Personal information about authors can be found throughout the entire collaborative content environment. Coordination features (e.g. no workflow support and only average project management support) are not among the strengths of this suite. However, a lot of additional cross-sectional and administrative functions depict Jive SBS as a technologically mature platform rated by Gartner (2009) as one of the market leaders.

PHProject got the best evaluation in the category coordination. This is no surprise as PHProject was designed as a groupware tool for project management. It offers standard project management features such as task and resource management as well as the coordination of schedules and the administration of meetings. PHProject also allows managing multiple projects. A weak spot of the system is little functionality in the category connection.

Alfresco Share was the most extensive tool for collaboration. Its specific strength is the administration of shared content. All features offered by Alfresco are highly integrated. Communication is the weak spot of the tool as email is not supported. Alfresco is an open-source E-Collaboration system and thus offers high adaptability. It is recommended for the collaborative work of small teams because for larger teams one quickly loses track.

Table 2 shows the results of the evaluation of Alfresco in the category collaboration. For each functionality Alfresco got the according score if the feature was supported (●) or not (○). In the future this score will be replaced by a more precise evaluation providing information not only if a feature is supported but in addition how well it is supported. This will be symbolised on a five level scale also using quarter, half or three-quarter circles (not included in Table 2 yet).

Table 2: Functionalities and weights of the category collaboration – scores for *Alfresco Share*.

Category Collaboration	Total Score 100,00	Supported	Alfresco Share Score 70,25
Shared Content Production	20,00		10,00
Wiki	10,00	●	10,00
Whiteboard	3,00	○	0,00
Synchronously Shared Documents	4,00	○	0,00
Shared Ideas / Brainstorming	3,00	○	0,00
Working together on the same objects	15,00		12,75
Social Tagging	6,75	●	6,75
Social Bookmarking	6,00	●	6,00
Social Cataloguing	2,25	○	0,00
Administration of shared content	50,00		47,50
Document Sharing	7,50	●	7,50
Image Sharing	2,00	●	2,00
Video/Audio Sharing	2,00	●	2,00
Restricted Access for Content	4,00	●	4,00
Restricted Access for Folder	2,50	●	2,50
Check in/Check out	5,00	●	5,00
Up & download	5,00	●	5,00
Versioning	4,00	●	4,00
Archiving	2,50	○	0,00
Folder / Shared Folder	7,50	●	7,50

Table 2: Functionalities and weights of the category collaboration – scores for *Alfresco Share* (Cont.).

Category Collaboration	Total Score 100,00	Supported	Alfresco Share Score 70,25
Shared Content Production	**20,00**		**10,00**
Content Tagging	4,00	●	4,00
Folder Tagging	1,50	●	1,50
Personal Site	2,50	●	2,50
Creating Documents out of the Shared Workspace	**15,00**		**0,00**
Textdocument	5,25	○	0,00
Calculation	1,50	○	0,00
Presentation	1,50	○	0,00
Graphics	0,75	○	0,00
…using MS Office	6,00	○	0,00

Figure 1 provides an overview of the assessment results for the five sample E-Collaboration systems in the basic categories depicting the strengths and weaknesses of these tools.
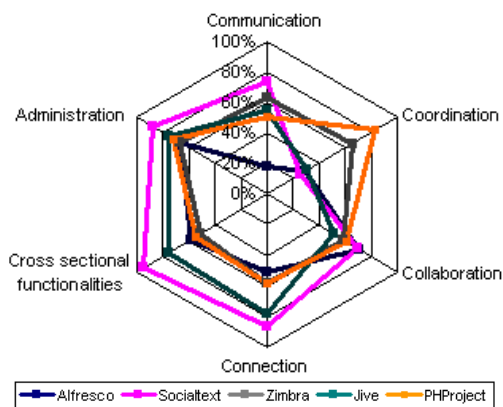
Figure 1: Overview of the assessment of five example E-Collaboration systems.

5 CONCLUSIONS

This market review presents a software evaluation approach adjusted to the special requirements of the steadily changing E-Collaboration market. Practitioners and IT-experts will find a set of evaluation criteria and an easy to adopt evaluation approach that provides an overview of the strengths and weaknesses of available software products. It can be easily adjusted to evaluate and select E-Collaboration systems for specific application scenarios. In this case, our set of features and

functional criteria provides an excellent basis that can easily be supplemented by technical criteria (e.g., concerning integration issues), financial criteria (e.g., licensing and distribution model) and vendor criteria (e.g., ability to execute).

This market review focuses on E-Collaboration systems that cover all aspects of social interaction. Therefore only software products have been tested that support all 4Cs of electronic collaboration: communication, coordination, collaboration and connection. The huge market of software products that cover only a fraction of functionality and provide specialised features in a smaller application segment has been excluded deliberately (for general market studies see e.g. Hinchcliffe, 2007; CMS Watch, 2009; Gartner, 2009). However, companies and institutions that want to implement or enhance a comprehensive electronic collaboration strategy will need to look at complete E-Collaboration packages.

Assessment of E-Collaboration systems based on the presented approach will continue on a continuous basis leading to a periodic report on the E-Collaboration marketplace. Ratings on the analysed software packages will be available in an online database in the future.

REFERENCES

AIIM (n.d.). What is Enterprise 2.0? Retrieved from http://www.aiim.org/What-is-Enterprise-2.0-E2.0.aspx on February 19, 2010.

Allen, C. (2004). Tracing the Evolution of Social Software. Retrieved from http://www.lifewithalacrity.com/2004/10/tracing_the_ evo.html on February 19, 2010.

Austin, T. and Burton, B. (2004). Define Collaboration Before Planning a Strategy. *Gartner Research*, Research Note G00124081.

CMS Watch (2009). The Enterprise Social Software & Collaboration Report 2009. Retrieved from http://www.cmswatch.com/Social/Report/ on February 19, 2010.

Cook, N. (2008). *Enterprise 2.0 - How Social Software Will Change the Future of Work*, Gower. Aldershot.

Gartner (2009). Magic Quadrant for Social Software in the Workplace. Retrieved from http://www.gartner.com/DisplayDocument?id=789812 respectively http://acquia.com/community/resources/library/2009-gartner-mq-social-software on February 19, 2010.

Hinchcliffe, D. (2007). The state of Enterprise 2.0, *Enterprise 2.0 Web Blog, ZDNet*. Retreived from http://blogs.zdnet.com/Hinchcliffe/?p=143 on February 19, 2010.

Illik, A. (2009). *Online Collaboration: Zusammenarbeit im globalisierten Team*, Ambit Informatik. Stuttgart.

Kock, N. (2005). What is E-Collaboration. In *International Journal of e-Collaboration*. Vol. 1, No. 1, p. i–vii.

Koplowitz, R. (2009). *The Forrester Wave™: Collaboration Platforms, Q3 2009*, Forrester Research

McAfee, A. (2006a). Enterprise 2.0: The Dawn of Emergent Collaboration. In *MIT Sloan Management Review*, Vol. 47, No. 3, p. 21–28. Retrieved from http://sloanreview.mit.edu/the-magazine/articles/2006/spring/47306/enterprise-the-dawn-of-emergent-collaboration/ on February 19, 2010.

McAfee, A. (2006b). Enterprise 2.0, version 2.0. In *Andrew McAfee's Blog - The Business Impact of IT* (May 27, 2006). Retrieved from http://andrewmcafee.org/2006/05/enterprise_20_version_20/ on February 19, 2010.

Riemer, K. (2007). The Market for E-Collaboration Systems - Identification of System Classes Using Cluster Analysis. In. Österle, H., Schelp, J., Winter, R. (eds.). *Proceedings of the Fifteenth European Conference on Information Systems*. p. 346-357.

Riemer, K. (2009). eCollaboration: Systeme, Anwendungen und aktuelle Entwicklungen. In *HMD Praxis der Wirtschaftsinformatik*. Vol. 267, June 2009, p. 7-17.

Teufel, S., Sauter, C., Mühlherr, T. and Bauknecht, K. (1995). *Computerunterstützung für die Gruppenarbeit*, Addison-Wesley.

Wilson, P. (1991). Computer Supported Cooperative Work: An Introduction, Kluwer Academic Pub.

A CASE STUDY OF THE E-MONEY APPLICATION IN JAPANESE PUBLIC TRANSPORTATION

Shoichi Morimoto

Department of Business Administration, Senshu University, 2-1-1 Higashi-mita, Tama-ku, 214-8580, Kanagawa, Japan
morimo@isc.senshu-u.ac.jp

Abstract: Japan leads the world in the field of a rechargeable contactless smart card used as a fare card of public transportation. The card triggered off the spread of Japanese e-money, however, the e-money situation has various intricate problems to tackle. Therefore, we have surveyed the spread process of the e-money and special circumstances of Japanese public transportation. In this paper we describe the business success factors and background. We also analyze and propose the solution against the problems and objectives for globalization of the market.

1 INTRODUCTION

Recently, the application of e-money has expanded into a means of payment of public transportation. The e-money for the fare is generally implemented by a smart card, particularly a contactless RFID smart card. Such a contactless smart card is used all over the world, e.g., Octopus card in Hong Kong, Suica card in Japan, Oyster card in London, Navigo card in Paris, T-money in Seoul, Compass Card in San Diego, Riocard in Rio de Janeiro, GoCard in Nigeria, and so on. Above all, Japan is an advanced country of the contactless smart card.

The spread process of Japanese e-money is very unique. Japanese people had not had interest in e-money at all, until the fare contactless smart card became practicable. After the card appeared in major rail services, e-money has spread rapidly, because a commuter can pass without stopping the very crowded ticket gate during the Japanese rush hours. The card can also be used at kiosks and vending machines inside stations. Moreover, recently the card has been able to be used at convenience stores, supermarkets, eating houses and the other shops outside stations. In addition, the bus trade has introduced the card. Thus various Japanese transportation companies have issued such a card with e-money. One innovative service with the card is coming out after another in quick succession and entry of the other trades into the market is increasing. Now the business of the contactless smart card with e-money in Japan has been a great success.

However, the e-money market has some problems to solve. Since each transportation company has issued its independent card, the Japanese transportation trade is flooded with many kinds (about 40) of contactless smart cards. The introduction of the card system is also very expensive, thus small companies which do not sufficiently have the capital strength (e.g., a local bus company) cannot even introduce the card and they are outdistanced. Moreover, FeliCa, the Japanese de facto standard of the contactless smart card, is not certified by the RFID international standard ISO/IEC 14443, that is, it is not global standard. The business of e-money in Japan is overly concentrated on the domestic market only. There is very little room for entry of overseas enterprises.

So far only the technological side of a contactless smart card and its system has been highlighted (Shiibashi, 2007). No studies have tried to survey the trade all over and to summarize the business success factors. Therefore, in this paper we survey the evolution process of e-money in Japanese public transportation and clarify the background and the problems which are caused by special circumstances of Japanese railways. We also discuss the secrets of success and solutions against the problems of compatibility and localization. The results in this paper open the Japanese e-money market and help its globalization. Moreover, they systematize the introduction of a contactless smart card into public transportation. Consequently, it will enliven the e-money market all over the world.

2 THE CARD SITUATION

2.1 Early History of Japanese e-Money

In 1998, the first Japanese e-money, VISA Cash, was introduced experimentally. However, this experiment has ended in failure. VISA Cash was implemented by a contact smart card. The user had to take the card from a holder and input the PIN code. Japanese people did not accept such troublesome operations.

In 2001, the first contactless smart card with e-money, named Edy, appeared. Edy is a pre-paid rechargeable contactless smart card, which uses Sony's FeliCa technology. FeliCa allows to send/receive data at high speed and with high security (Kurosawa et al., 2003). It does not need a battery to operate. Further details of the FeliCa technology will be presented later. Edy was able to be used only at convenience stores then. Japanese people did not particularly interest in Edy yet, because e-money did not have a good image at the time owing to the former e-moneys.

In the same year, the Suica service started in the metropolitan area (Shirakawa and Shiibashi, 2003). Suica using the FeliCa technology is a rechargeable contactless smart card used as a fare card on train lines. Suica was able to be used only as the fare payment in the limited lines then; nevertheless it had a circulation of one million only in 19 days and two million only in two months. The service has been expanded besides the fare payment after 2004 successively.

The e-money trade has been an arena of rival cards since Suica achieved a great success. Many public transportation companies followed the service and issued similar cards. We describe details of such cards and their development in the following subsections.

2.2 The Cards in Public Railways

Japanese railways are classified into two types; JR (Japan Railways) or non-JR. JR used to be the national railway JNR (Japan National Railways), which dissolved and separated into seven (six passenger railways and one freightage) JR companies in 1987. The passenger railways are separated by region. Figure 1 shows each JR company's area.

Although Japan is a small island country, the characteristic of the people of each region varies according to the locality. For instance, consumer behavior and interest differ from the Kanto region (Tokyo) to the Kansai region (Osaka/Kyoto) very much. Each JR company has to work out management policies which bear closely on the needs of local people. Each JR
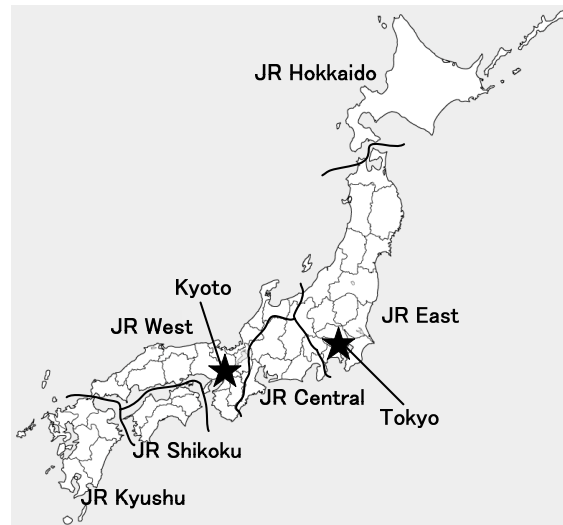
Figure 1: The region of JR companies.

company also considers the other JR companies competitors which scramble for business chances. Therefore, the companies developed their own cards shown in Table 1. All the cards are implemented by FeliCa.

Table 1: The contactless smart cards of JR group.

Company	Card name	Number of stations
JR Hokkaido	Kitaca	465
JR East	Suica	1705
JR Central	TOICA	404
JR West	ICOCA	1222
JR Shikoku	–	259
JR Kyushu	SUGOCA	560

A user can charge money on to the card at ticket vending machines inside each station. The charged money is a profit of the JR company in the card region. Thus, at first each card was not able to use in the other regions. The introduction of contactless smart cards to the Shinkansen line, which is a high-speed railway line, had also been put off for some years. The Tokaido Shinkansen line is the main artery crossing the Japanese mainland. JR Central has jurisdiction over the line. Because the line lies across the regions, JR Central and the other JR companies have been at odds with each other over the profit at the stations in the Tokaido Shinkansen line.

The specification of each card system is also different from the others. For example, the Suica system closes a ticket gate if e-money on the card is under the starting fare when a passenger enters. Meanwhile the ICOCA system does not close it unless the e-money is empty, since the Kansai region people generally tend to dislike a time-consuming operation. TOICA accepts the empty card. Moreover, the fare adjustment

rule of each JR company is different. The difference is also caused by the introduction order. The JR companies following the formers can watch the response to the card. The year of each introduction is shown in Figure 2 (the number in parenthesis). The compatibility is implemented by a software program in the card reader, however, each card system must have the programs against all the other cards.

2.3 The Cards in Private Railways

The private railways (non-JR railways) also started the fare contactless smart card service. The services are classified into a joint capital type or an independent type. In the metropolitan area, the private railways established a new association PASMO Co. Ltd. of the contactless smart card business by joint capital. The PASMO card is perfectly compatible with Suica. Therefore, transfer of JR lines and private/underground railways has been very convenient; we can change one line to another only with one contactless smart card in the Kanto region. It was the epochal event of Japanese railways history, because JR and non-JR had been always competitors one another. Similarly, the Kansai region private railways have issued the common card PiTaPa. PiTaPa is a novel card in Japanese public transportation, which is a postpay type card like a credit card for the impatient Kansai region people. PiTaPa also implements various discount services. The private railways in the other regions have also issued their own cards shown in Table 2.
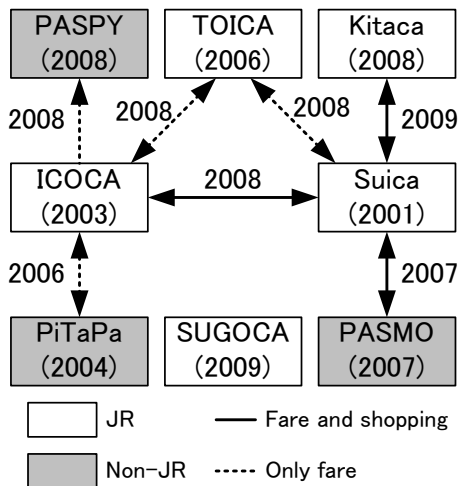
Figure 2: The compatibility of the cards.

The private railways cards have the problem of compatibility. Each card is compatible only with the JR card in its same region (see Figure 2). For example, the PASMO card can use only in the JR East

Table 2: The contactless smart cards of the private railways.

Region	Card name	Number of members
Kanto	PASMO	103 (77)
Kansai	PiTaPa	37 (17)
Hokuriku	ICa	4 (3)
Chugoku	PASPY	13 (9)
Shikoku	DESUCA	5 (4)
Kyushu	RaPiCa	4 (3)

(Suica) region and is not compatible with any other cards. Thus the member companies which have fully the capital strength want to cooperate with the other regions; on the other hand the small member companies oppose it. Each member company has to pay a membership fee and maintenance costs of several hundreds millions Yen (about several millions USD) every year. Moreover, it must pay several millions Yen whenever changes of the system happen. The small or local companies do not interest in use in the other regions, because they cannot make a profit on the mutual use any more.

Besides the cards in Table 2, some local private railways had to issue their independent card for lack of an alternative. Such cards do not have any compatibility.

2.4 Bus Companies

The bus companies have joined the association of the private railways. The number in parenthesis in Table 2 shows the number of bus member companies out of the total. The bus companies' case also has the same financial problem as well as the private railways case.

The bus member companies of PASMO have about 15,000 buses in all, however, the contactless smart card system has been introduced into only 30% of the buses. Because besides the membership fee and maintenance costs, the introduction of the system requires from seven to eight hundred thousand Yen (from eight to nine thousand USD) per bus. Nowadays Japanese people think it a matter of course that the contactless smart card can be used in public transportation. The bus companies which have not introduced it may lose customers.

3 THE CARD TECHNOLOGY

In this section, we describe why the card has spread rapidly from the technology viewpoint.

3.1 Why Contactless?

Japanese railways have special circumstances which other countries do not have (and cannot understand). Almost all Japanese people in the metropolitan areas take the train daily to work. Therefore, a station and a train are terribly crowded in rush hours. About 60 commuters per minute pass a ticket gate in peak hours. In Europe, it is 30 per minute. Additionally, Japanese train service is strictly on time. The machine of Japanese ticket gate requires high throughput.

There are many kinds of train tickets in Japanese railways. The ticket gate must examine varied train tickets. The gate also does not care insertion direction and the front/back of a ticket. One ticket gate accepts both of entry and exit of passengers (Figure 3). The Japanese ticket gate is a very complex machine. Therefore, the ticket gate machine tends to break down and its maintenance cost comes high. If the physical contact can be avoided as much as possible, the maintenance cost is reducible.

Consequently, a contactless smart card has been suitable against the foregoing circumstances. The first and most successful card Suica of JR East has succeeded in cutting down the maintenance cost by 1/3.

Figure 3: The ticket gate machine.

3.2 FeliCa Technology

Almost all Japanese contactless smart cards are implemented by FeliCa. To solve the congestion mentioned above, JR East invited tenders for the card system which can examine one card in 0.1 second. That is, it has to communicate at 211.875 kbit/s. It was very high level of specification. Moreover, JR East requested that the card reader should examine two or more cards in piles. Only FeliCa had satisfied these requirements. Since FeliCa had been already introduced into public transportation in Hong Kong, it was also reliable (Chau and Poon, 2003). After JR East adopted FeliCa, it has become the de facto standard.
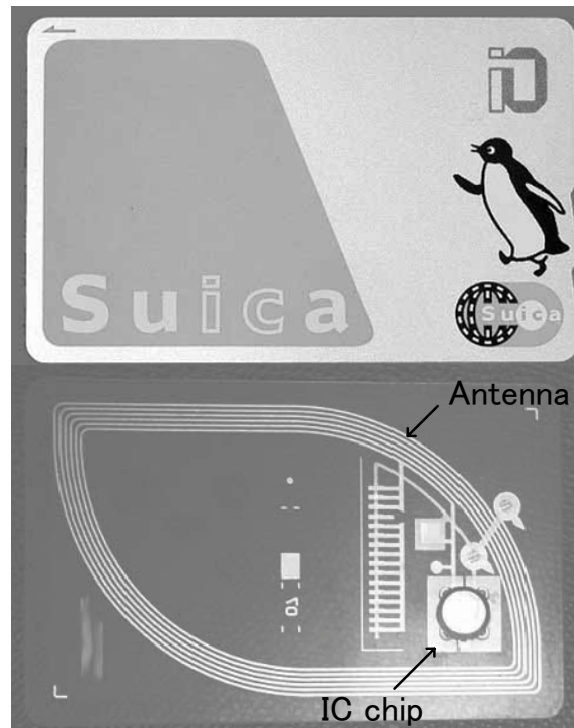
Figure 4: The Suica card and its cutaway drawing.

FeliCa uses Manchester coding at 212 kbit/s in the 13.56 MHz range. It can also operate without a battery using the electromagnetic induction. Thus it is thin and light in weight. The coil like a leaf shown in Figure 4 is an antenna. The leaf form is a device for reading two or more cards. When the antenna comes close to the reader, an electric current is generated. FeliCa reads and writes data with the power. It can implement various applications, if only software in the reader is rewritten. Moreover, the security of FeliCa has been certified by ISO/IEC 15408. These FeliCa functions are suitable for the smart card of public transportation.

However, FeliCa has been rejected by the RFID international standard (ISO/IEC 14443, 2008). It has been certified by the NFC (Near Field Communication) international standard later (ISO/IEC 18092, 2004). ISO/IEC 14443 provides two types of the RFID card. The type A uses Modified Miller coding at 106 kbit/s, which is developed by Philips in Netherlands. The type B uses NRZ (Non Return to Zero) coding at 106 kbit/s, which is developed by Motorola in the U.S. The type A is most widely used in the world. Nevertheless, the type A and B are hardly used in Japan.

4 DISCUSSION

Nowadays, 50% of Japanese people in the major cities have more than one contactless smart card with e-money. Suica certainly triggered the spread of e-money. Japanese people have understood convenience of e-money by Suica. The graph shown in Figure 5 backs up this fact. The Edy e-money only for shopping has been gradually increasing, since Suica appeared. The spread of the mutual use among
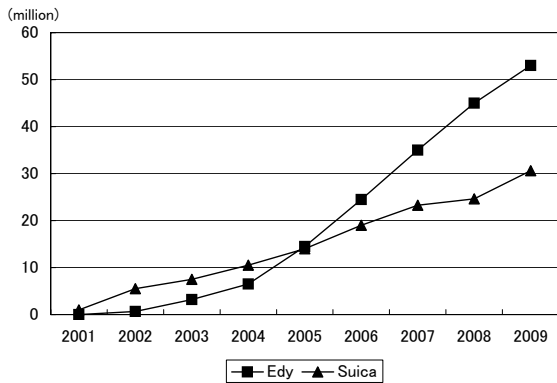
Figure 5: The number of Edy and Suica cards.

the cards will promote competition among the major companies and improve the quality of the service. On the other hand, it may merely encourage the disparity between local areas and metropolitan areas. Japan is a conspicuous aging society. Therefore, the local transportation companies aim to revitalize depopulated areas using e-money with the contactless smart cards. However, the mutual use has few merits for the local transportation companies. In addition, it requires the software development, but the local transportation companies do not have software engineers. They have to request a software vendor to develop it: the development cost is about one million Yen (10,000 USD). Now the government partially supports the local companies only in the initial costs. It should support them further.
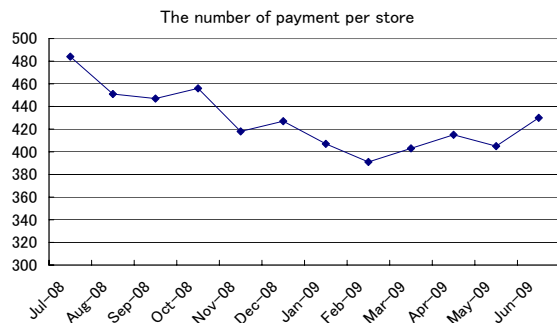
Figure 6: The number of payment per store with Suica.
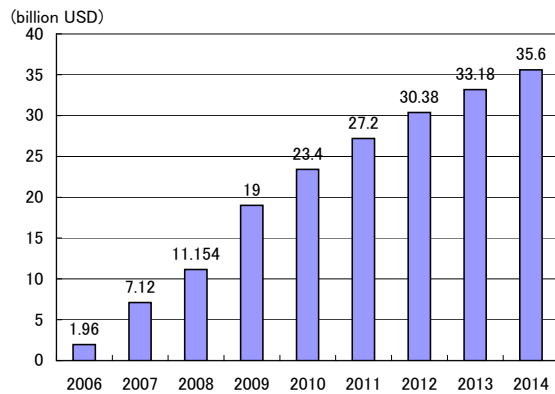
Figure 7: The prospect of the e-money market in Japan.

5 CONCLUDING REMARKS

The number of the contactless smart cards in Japan has virtually reached the ceiling; nevertheless the number of the payment per store is gradually decreasing (Figure 6). For the further spread, each company or association has to increase member store inside and outside stations and to propose novel and attractive privileges. JR East has made full use of the advantage in the metropolitan area. Suica money is mutually exchangeable for the mileage point of ANA (All Nippon Airways). Suica is also developing novel services using data-mining from the use records as lifelogs.

The contactless smart cards with e-money have spread in the world. For instance, PayPass of Master-Card has a circulation of 60 millions. Sooner or later Japan must open the contactless smart card market. Before opening the market, the Japanese e-money trade has to solve the problem of the mutual use among FeliCa, the type A and the type B of ISO/IEC 14443. ISO/IEC 18092 standardizes NFC, which has unified FeliCa and type A. Similarly, type B has been added to NFC (ISO/IEC 21481, 2005). The card reader adapting to NFC can use FeliCa, type A and type B. Such reader is spreading; however, NFC is upward compatible for the hardware protocol. The system must have software applications for every chips. Thus, NFC cannot become a radical solution. Moreover, FeliCa has a serious problem. Generally, a recent mobile phone has a contactless smart card built-in. In response to the situation, GSMA (Global System for Mobile communications Association, 2010) has standardized a contactless smart card system for a mobile phone and adopted type A and type B as a global standard. All over the world, e-money services with such mobile phones have already begun. Since Japanese mobile phones, which are equipped with FeliCa, have the protocol stack different from the mobile

phones with type A/B, Japan may be isolated in the mobile field too. In the future, Japanese mobile phone should have not only FeliCa but the other global IC chip and use different applications for every chips.

Nomura Research Institute showed very interesting statistics shown in Figure 7 (Nomura Research Institute, 2008). They expect that the e-money market in Japan will be expanded further. Thus there are many business chances for the domestic and foreign companies. The foreign company which entries into the Japanese market must understand the national character and special circumstances of Japanese public transportation. The architecture of the payment system with contactless smart cards in Japanese public transportation has been referred to the IFMS (Interoperable Fare Management System) standard (ISO/IEC 24014, 2007). Japan has been nominated for the editor of ISO/IEC 24014 Part 2: recommend business practice for set of rules. The organization is currently summarizing the business model as the technical report. This standardization will be certainly helpful for the spread of the e-money application of public transportation in the world.

REFERENCES

Chau, P. and Poon, S. (2003). Octopus: An e-cash payment system success story. *Communications of the ACM*, 46(9):129–133.

Global System for Mobile communications Association. (2010). Gsm world –connecting the world–. http://www.gsmworld.com/.

ISO/IEC 14443 (2008). *Identification cards - Contactless integrated circuit(s) cards - Proximity cards -*.

ISO/IEC 18092 (2004). *Information technology – Telecommunications and information exchange between systems – Near Field Communication – Interface and Protocol (NFCIP-1)*.

ISO/IEC 21481 (2005). *Information technology – Telecommunications and information exchange between systems – Near Field Communication Interface and Protocol -2 (NFCIP-2)*.

ISO/IEC 24014 (2007). *Public transport – Interoperable fare management system – Part 1: Architecture*.

Kurosawa, A., Morita, T., and Kusakabe, S. (2003). Contactless ic card technology "felica" and new approach. In *Proceedings of International Symposium on Seed-up and Service Technology for Railway and Maglev Systems*, pages 323–327.

Nomura Research Institute, Ltd. (2008). Survey on e-money reveals rapid expansion of retail e-money cards, showing signs of growing competition for becoming the primary e-money card. http://www.nri.co.jp/english/news/2008/080717.html.

Shiibashi, A. (2007). High-speed processing and high-reliability technology in an integrated fixed-line and wireless, autonomous decentralized ic card system. *Systems and Computers in Japan*, 38(9):1–10.

Shirakawa, Y. and Shiibashi, A. (2003). Jr east contact-less ic card automatic fare collection system 'suica'. *IEICE Transactions on Information and Systems*, E86-D(10):2070–2076.

CAWP
A Combinatorial Auction Web Platform

İbrahim Cereci and Hürevren Kılıç

Computer Engineering Department, Atılım University, İncek, Ankara, Turkey
{icereci, hurevren}@atilim.edu.tr

Abstract: Online auctions, including online Combinatorial Auctions, are important examples of e-commerce applications. In this paper, a Combinatorial Auction Web Platform (CAWP) is introduced. The platform enables both product selling and buying capabilities that can be realized in a combinatorial way. CAWP supports a Sealed-Bid Single-Unit type of Combinatorial Auctions. Easy customization for any selected problem domain is a distinguished feature of CAWP. Platform users are not expected to have any technical knowledge about how to solve the Winner Determination Problem (WDP) known to be critical for profit maximization of the auctioneers in Combinatorial Auctions.

1   INTRODUCTION

Auction is a trading process where auctioneer provides goods or services and buyer bids to these goods or services. At the end, the highest bidder wins. Online auctions are the auctions which are held over the internet. Rapid growth of internet makes online auctioning important, as it reduces the time and space cost of the offline auctioning mechanisms. Combinatorial Auctions (CA) are the auctions that bidders place bids on combinations of items rather than a single item (Vries and Vohra, 2010). CAs are commonly used in application areas like transportation (Kwon et.al., 2005), bus routes, airport landing rights, power exchanges, carbon permits, and radio spectrum for wireless communications services (Milgrom, 2000) For example, in 2002, Nigeria sold regional fixed wireless access licenses on a sealed-bid combinatorial auction (Koboldt et.al., 2003). Similarly, allocation of web services via CAs is possible. The requirement of more than one web services to be elaborated at the same time implies a form of bidding that supports web services combinations (Lin et.al., 2008). CAs also used in Supply Chain Formation. The supply chain formation demands difficult coordination issues for distributed negotiation of the protocols to be solved. Parties must negotiate for multi level production relationships with important interdependencies among inputs and outputs of each level. CAs

addresses this problem by global optimization over expressed offers to engage in compound exchanges (Walsh et.al., 2000).

In practice, CAs are popular because they give bidders a capability to express their complete preferences. Especially, if the items in the auction are complementary, set of items may be valued as more than the sum of values for each individual item. At the same time, the auctioneer may obtain higher benefit by initiating a CA instance. This is because of allowing bidders to express their preferences in combinatorial way, which may results in better auction revenues (Cramton et.al., 2007). Automation of the CA is clearly important, because sellers may want to maximize their revenues and let their bidders flexibly express their preferences while bidding for the items in the auction over the internet.

One of the main problems for the auctioneer in an online CA is to decide about which bid(s) will be allocated (or chosen as the winner(s)). CA allows bidder to bid bundles of items in an auction while these bundles may overlap. The aim is to find a subset of all given bids that will maximize the resulting revenue of the seller. In literature, this problem is called as Winner Determination Problem (WDP) known to be NP-complete (Gottlob and Greco, 2007). Online CAs cannot perform well for the unbounded large scale problems. But with giving limitation to maximum number of items in an auction, problem size can be reduced to a solvable instances. To the best of author's knowledge, there

is no CA platform realizing Consumer to Consumer (C2C) auctions. In this paper, an online configurable Combinatorial Auction Web Platform (CAWP) is introduced (Cereci, 2009). The platform can be used by consumers who want to sell or buy goods in a combinatorial way. Basically, the consumers are not required to know the details about how to solve the WDP. CAWP is developed by using open source tools and technologies. It is operating system independent. The sellers and buyers can interact with the system through a simple web browser without any additional program installation.

The rest of the paper has the following organization. In Section 2, a background information about general auctions and their automation are provided. In Section 3, combinatorial auctions and the winner determination problem are explained. Also, two alternative solutions to WDP are discussed. In Section 4, technical details, performance results and an example usage of CAWP are given. The last section includes the conclusion and future works.

2 ONLINE AUCTIONS

Emergence of the Internet has changed the way people buy and sell goods. New types of electronic marketplaces have been developed to create more efficient markets (Bakos, 1998). Online Auctions have been one of the most successful electronic markets (Wolfram|Alpha, 2009). Success of the online auctions comes from the capabilities that they provide both to buyers and sellers. As a buyer, one can bid on large number of items from different sources and he/she has the potential to find goods in lower prices. As a seller, you can reach great number of potential buyers.

2.1 Auctions

There are four common auction types. Most of the other auction types are derived from these basic four auction types (Klemperer, 2004).

1. *First-Price Sealed Bid Auction:* All bidders submit their valuations in sealed bids, simultaneously. By this way it is guaranteed that no bidder knows the bid of the others. The highest bidder who pays the price gets the good.

2. *English Auction* (a.k.a. *open-cry ascending price* auction): English auction is the most common form of auction used today. In this form, bidders bid openly against each other. Each bid should

be higher than the previous one. The auction ends when no bidder is willing to raise the final bid or bidding period is over. The highest bidder gets the good.

3. *Dutch Auction* (a.k.a. *open-cry descending price* auction): This auction type is similar to the English auction. The auctioneer begins with setting a high price to the good. Initial price is gradually lowered until one bidder accepts to pay that amount. Last announced amount is paid by the bidder.

4. *Vickrey Auction* (a.k.a. *sealed-bid second-price* auction): It is identical to the first-price sealed-bid auction except that winning bidder pays not his bid amount but the second highest bidder's amount.

2.2 Electronic Auctions

Electronic auctions became an important part of the electronic trading. In general, complete trading process of any online auction has the following steps (Kumar and Feldman, 1998):

1. *Initial buyer and seller registration:* All parties are authenticated.

2. *Setting up a particular auction event:* Goods are described; auction rules are set and auction is started.

3. *Scheduling and advertising:* Upcoming auctions are notified to attract potential buyers.

4. *Bidding:* All the bids are collected. Bid validity is verified during bidding period and bids are placed until the bidding period is over.

5. *Evaluation of bids, closing the auction:* The auction closing rules are applied and the winner bids are determined. Winners and losers are notified back.

6. *Trade settlement:* Payment and good delivery are realized.

Furthermore, for the sake of standardization, every electronic auction platform are also expected to support the following properties (Omote, 2002):

1. *Anonymity:* Loser bidders should not be identifiable.

2. *Non-cancelability:* A winner is always identified that he cannot deny having bid to the auction.

3. *Public verifiability:* Anybody could publicly verify the winning bid is really the highest value and valid.

4. *Unforgeability:* Impersonation of sellers and bidders should be prevented.

5. *Robustness:* Auction process should not be interrupted, even due to invalid bids.

6. *Fairness:* Every bid should have the same priority; there should be no favor to any individual's bids.

7. *Efficiency of bidding:* The computation of determining a winner bid and verifying that should be practical.

In CAWP, most of the above processing steps are realized together with the mentioned properties. The details of them will be given in subsection 4.2.

3 COMBINATORIAL AUCTIONS

CAs can be categorized according to certain criteria described below:

1. Categorization based on bidding style:

- *Open-Bid Combinatorial Auctions:* Bidders are aware of competing bidders' bids. All bids are publicly announced.

- *Sealed-Bid Combinatorial Auctions:* Each bidder is only aware of his/her bids. After bidding process is completed and winner is determined, it can be announced. Hiding the bids is necessary during bidding time.

2. Categorization based on the number of goods:

- *Single-Unit Combinatorial Auctions:* Amount of the each individual item is one. For example, if there are five identical items, they must be placed to the auction as different items which have the same product information.

- *Multi-Unit Combinatorial Auctions:* Amount of the individual items may be more than one.

3. Categorization based on pricing:

- *Reserve Combinatorial Auctions:* Seller may put a base acceptance price on each item during auctioning. Since in CA bundles of items are bid together, amount of bid should be more than the sum of the items' base prices in the bundle.

- *Non-Reserve Combinatorial Auctions:* Seller cannot put a base acceptance price to their items in the auction. Winner(s) pay the amount they bid and get the items even they are below their original value.

- *Reverse Combinatorial Auctions:* Buyers may want series of items and sellers bid group of that items. Least expensive bids are accepted.

CAWP implements a *Sealed-Bid Single-Unit Combinatorial Auction* mechanism where items can have reserved prices. They are set by the Auctioneer. Every item put in an auction is single unit, if two identical items is needed to be put in an auction, they should be placed separately.

In CAs, bidders are allowed to express themselves freely and place any combination of bid items for the auction. However, this comes with an explosion of the size of the solution space. Winner Determination Problem is the problem of deciding the allocation of winner bids, in a set of bids placed to the auction, so that the revenue of the auctioneer can be maximized.

Formally, let I be the set of items under consideration and R^+ be the set of non-negative real numbers. Then, we say that a bid $b = (I_b, P_b,)$ is an element of $S = (2^I - \{\emptyset\}) \times R^+$. That means any subset of power set of items I other than the empty set may have an assigned value decided by its bidder. Let B be a subset of S. A set $F \subseteq B$ is said to be *feasible* if $\forall b,c \in F$, $c \neq b$ and $I_b \cap I_c = \emptyset$. That is no two items in bidding subsets are the same. Also, let $\Phi(B)$ be the set of *all possible feasible allocations* for B. Further, let $I(B) = \bigcup_{b \in B} I_b$ be the set of goods contained in the bids of B.

Definition 1: *Winner Determination Problem* is to find an allocation $W \in \Phi(B)$ such that $\forall F \in \Phi(B)$ the following should hold

$$\sum_{b \in F} P_b \leq \sum_{b \in W} P_b$$

Such allocation is said to be optimal or revenue maximizing (Brown et.al., 1999).

WDP is hard because one would need to check all subset of the bids to identify whether they are feasible (no conflict of items) and how much revenue they may provide. A feasible subset of the bids that has the maximum revenue is the optimal solution. There are 2^k subsets of bids where k being the number of bids (Cramton et.al, 2006).

In general, there are three main factors affecting the solution time for a given WDP instance. These are the number of goods, number of bids and distribution of the bids. If there are some *dominant bids* in the system, solution can be found in dramatically shorter time. This is because when a solver accepts a dominant bid, it helps maximizing the auctioneer revenue and reduces the solution space, causing a solution to be found faster.

Combinatorial Auction Structured Search (CASS) (Brown et.al., 1999) and Combinatorial Auction Branch on Bids (CABOB) (Sandholm et.al., 2001) are two known algorithms for efficient solution of WDP:

1. *Combinatorial Auction Structured Search (CASS) Algorithm:* CASS uses exhaustive search for determining optimal solution. It suggests a simple brute-force search approach supported by four

significant heuristic improvements. CASS structures (or shapes) the search space in order to avoid conflicting bids with some overlapping items. It keeps the result of the searches done up to a point and prunes the search tree. Together with CASS, a test suite is developed to create sample auction setups and to test the systems performance. This test suite is known as CATS (Cramton et.al., 2007)(Brown et.al., 2000).

2. *Combinatorial Auction Branch on Bids (CABOB) algorithm:* CABOB is a tree search algorithm where tree is branched on bids. It makes a dept-first Branch&Bound search on the tree. Branch&Bound search provides a systematic enumeration of all possible solutions and it prunes large subsets of fruitless candidate solutions by using upper and lower bounding (Land and Doig, 1960). In CAWP, we prefered to use a solver that supports the CASS algorithm.

4 COMBINATORIAL AUCTION WEB PLATFORM

In this section CAWP will be introduced, with its functionalities according to the system and the user requirements. Technical details and performance results and a comparison with another CA system are also given.

4.1 General Properties of CAWP

Among the six steps describing complete trading process (given in subsection 2.2), CAWP fully supports the steps 1, 2, 4 and 5. Related to step 3, there is no advertisement support but the users can see and join directly to (or just observe) the product sets open to bidding. Also, there is neither payment nor good delivery tracking capabilities of the system. The requirements for a typical online auction system has also been given in subsection 2.2. Most of these requirements are satisfied by CAWP as described below:

Loser bids are not made known to all bidders (*anonymity*). Bidder logs in to the system before bidding so a winner is always identified (*non-cancelability*). Public verification of the winning bid is really the highest value is not possible. This is because the loser bids are not announced publicly. Note that even this could be the case; such verification would require re-solving the WDP under consideration. The only mechanism against impersonation is the username/password usage.

Impersonation of another bidder is not possible without stealing their user name and password (*unforgeability*). A person can impersonate another person in the CAWP, and jeopardize the fairness of the trade, but after single act this person can be notified and receive bad comments, or it can go to account suspension. CAWP prevents bidders to place invalid bids, and keeps the auction process uninterrupted (*robustness*). In CAWP bids or bidders have no priority (*fairness*). Finally, because of the algorithmic complexity of the WDP problem, for some problem instances the computation of determining and verifying a winner bid may not be practical (*efficiency of bidding*). In CAWP, this problem is tried to be handled by putting "at most 30 item per auction" rule for sellers. By this way, one can get a response from the WDP solver in an acceptable time.

Below is the list of implemented CAWP properties based on system/user requirements that are considered during development:

- Authentication of the users is a must for both bidding and creating auctions.
- Any internet user can enter to the CAWP site and view the products information and auctions. However, user must be logged in to get bidding and auctioning capabilities.
- There is no hierarchy or priority between users of the system.
- The system provides its auctioneers to set some parameters of the auction including auction name, auction end time, items in the auction, items' base prices.
- Bidders can bid any bundle of items as long as they are in the same auction.
- The system assists bidders with the minimum acceptable amount of bid being the sum of base prices of items in the bid.
- Users can withdraw their bids until bidding session is over.
- Auctioneer can drop an auction if there are no bids placed on any goods in the auction.
- Buyers and sellers can write comments about the people they trade with.
- Platform has an internal messaging system. Sellers and buyers can send private messages to each other.
- Number of items that can be on a single auction is maximum 30.
- Real procurement of the goods is realized between seller and buyer. CAWP only gives buyers and sellers capability to comment about their actual trading experiences.

4.2 Technical Details and Performance Results

Technically, CAWP is constituted from four main components. The first one is the *web-component* using which users can create auctions, and make bid. The winner determination problem is generated and results are notified via this component. Second component is the *solver-component* that gets the problems generated by the *web-component*, solves them and returns the solution. Third component is the *database-component* which is used by the web and solver components. Bid, auction, and bidder information are all kept by *database-component*. *Solver-component* gets necessary data from the *database-component* to create the winner determination problem file. After the solver's execution, results are put back into the database. The last component is the *customization- component*. This component is necessary to enable CAWP without knowing web programming. Site customization can be achieved by using this component. Figure 1 shows how WDP solver program interacts with the remaining components of the platform. *Optimizer-handler* seen in Figure 4 aims to provide interoperation of CAWP web platform and the CASS solver. Throughout the process, the handler plays the main *control-unit* role. It first checks whether there exists a WDP to be solved or not. If any, it translates the problem description kept in the database to the acceptable format of the CASS solver. Concurrently, it calls the CASS solver. Consequently, the problem is solved, the result is taken as generated output file, and put back into the database.

All technologies used in CAWP are open source. The overall system performance clearly depends on the performance of the technologies used. Performance issues related to the database system and web server that are used can be found in (MySQL, 2005) and (Apache, 2009), respectively. Aside from these, the critical component that effects the general performance is clearly the WDP solver. The relative performance of the CASS solver has been evaluated in (Brown and Shoham, 2009). In order to be able to better tuning of the CAWP system, we executed a series of performance test scenarios on Intel® Core™2 Duo 2.66 GHz CPU, 4GB Ram, Windows XP Professional environment. We created 10 different scenarios for every different number of good and bid instances. The test scenarios are generated with CATS tool (Brown et.al., 2000).Table 1 shows the average completion time of the executions in seconds. Even if there are 5000

bids to a single auction, the solver can still produce an answer within minutes. Also, if the number of bids is small even the auctions with greater number goods can be solved quickly. But still the major parameter that effects the solution time is the number of goods.
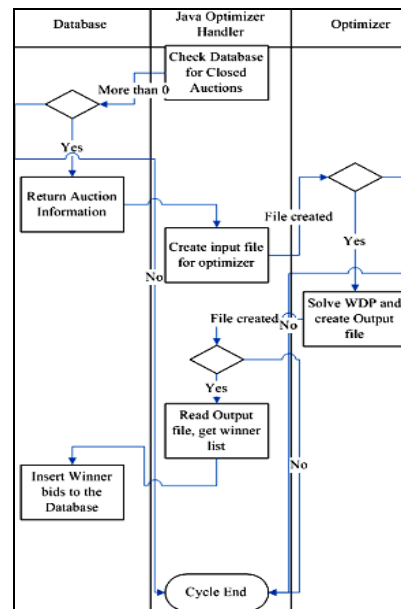
Figure 1: WDP solver interaction schema.

When we keep the number of bids as 5000 and increase the number of goods from 30 to 50 the completion time of the problem increases from 151.860 seconds to 3713.246 seconds, dramatically.

Increasing number of concurrent CASS solvers running on a single machine reduces the overall performance of the system (see Table 2). If aim is to increase the throughput of the system, concurrent solvers should run on different machines. One solver per processor gives the highest expected performance. The results in Table 2 are obtained by taking runs on a dual core machine where two or more solvers run on different CPUs.

In Figure 2, the NPC aspect of the problem solving operation can be seen. The completion time is sensitive to the number of goods variable *I* especially when *I* reaches to 50.Note that the CAWP system is not evaluated by their users in terms of its usability, yet.

Table 1: Average completion time results for the system tests.

Number of Goods	Number of Bids	Results (in seconds)
10	100	0.002
10	1000	0.016
10	3000	0.210
30	100	0.003
30	1000	0.047
30	3000	12.840
30	5000	151.860
40	100	0.015
40	1000	6.552
40	3000	126.017
40	5000	177.721
50	100	0.020
50	1000	8.424
50	3000	3713.246

Table 2: Performance results for many solvers running on a two- processor machine.

# of Solvers	# of Goods	# of Bids	Results (in seconds)
2	30	3000	17.34
3	30	3000	23.87
4	30	3000	35.31
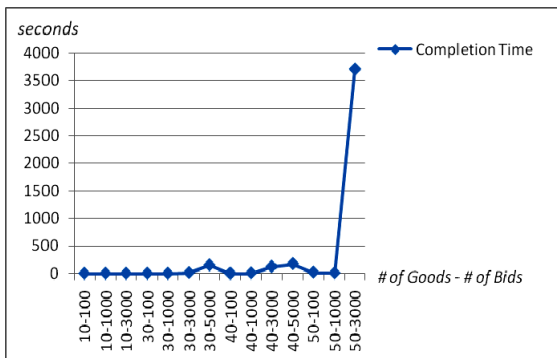
Figure 2: Completion time results for different instances of number of goods and number of bids pairs.

Finally, for our comparative purpose, we considered an example CA system namely Online Iterative Combinatorial Auction System (OICAS) (Fang and Wang, 2005). It is a prototype system implemented using Visual Basic and Microsoft Access. OICAS has following characteristics:

- The auctioneer can determine who can participate in an auction.
- Only legitimate users (bidders) can participate in an auction.
- Bidders can bid for any bundle of items with acceptable price.

- The system tells bidders the required minimum winning price for their bids automatically.

The main difference between OICAS and CAWP is that OICAS is an *offline* prototype implemented in Microsoft Visual Basic language and Microsoft Access database. On the other hand, CAWP is a true web application implemented using PHP for server side scripting, MySql for database management, Java for implementing optimizer handler, and C for the WDP solver. These components can be deployed and run on different machines. Also, although OICAS permits its auctioneers to choose their bidders, CAWP believes in the fairness for the bidders and it does not allow an auction to be limited for only a certain group of bidders.

4.3 An Example CAWP Usage For Custom Built Furniture Combinatorial Auctioning

A carpenter producing custom built furnitures can sell combinations of his products via CAWP in order to increase his revenue. However he needs a technical assistance to deploy and public his furniture combinatorial auction website, if otherwise he can do it himself. As the first step each different product categories are introduced to the system. Then an auctioneer account for the carpenter should be created. A list of available products to be auctioned are entered to the system by giving product name, category, description, picture, serial number and reserve price. Following this, an auction instance including auction name, duration and list of target items is generated. The auctioneer can open more than one auction. After this the bidding period is started. Within the auction period the system accepts bids from its registered users. For each bid the bidders should give the list of items to bid on and a valid price. By the end of auction period the system automatically initiates the optimizer in order to solve the generated WDP problem. The bidding results are announced at the site in individual basis. In other words a bidder is only informed about his winning status, rather then the others. For the time being the rest of the trading process is not supported by the system. On the other hand the auctioneer or the bidders can enter comments about their experience with the trading process.

5 CONCLUSIONS

In this paper, an online combinatorial auctioning platform CAWP is introduced. The platform provides its users to create and to participate in combinatorial auctions without having to care about either the complexity of the WDP or its efficient solution. The performance requirement of the system is clearly more higher than a typical online auctioning system. This is mainly due to the required involvement with an NP-Complete WDP problem. In our solution different technologies combined together and integrated under the platform. The technologies include server side web scripting, database management, solver handler, and the WDP solver. All these technologies, except the WDP solver have been created, originally. Using open source technologies enabled us to build an operating system independent platform.

In future, CAWP is planned to be supported by a third party payment system in order to achieve better trading opportunities. Also, the system may support a realistic mechanism to reward and penalize its users. A better WDP solver or a general purpose solver package can be adapted to the system in order to still increase the WDP solution performance.

REFERENCES

Vries S., Vohra R. 2010. Combinatorial Auctions: A Survey. http://www.cis.upenn.edu/~mkearns/teaching/cgt/combinatorial-auctions-survey.pdf. As of April 2010

Kwon R.H., Lee C., Ma Z., 2005. An Integrated Combinatorial Auction Mechanism for Truckload Transportation Procurement. In: Technical Report, Mechanical and Industrial Engineering, the University of Toronto.

Milgrom P., 2000. Putting Auction Theory to Work: The Simultaneous Ascending Auction. In: Journal of Political Economy, Vol. 108, No. 2.

Koboldt C., Maldoom D., Marsden R., 2003. The First Combinatorial Spectrum Auction. In: DotEcon DP http://www.dotecon.com/publications/dp0301.pdf. As of April 2010

Lin S., Chen B., Liu C., Soo W., 2008. Web Service Allocations Based on Combinatorial Auctions and Market-based Mechanisms. In: *Computer Supported Cooperative Work in Design (CSCWD 2008)*, vol., iss. Page(s):452 - 458

Walsh W.E., Wellman M., Ygge F., 2000. Combinatorial Auctions for Supply Chain Formation. In: ACM Conference on Electronic Commerce.

Cramton P., Shoham Y., Steinberg R., 2007. An Overview of Combinatorial Auctions. In: ACM SIGecom Exchanges, Vol. 7, No. 1.

Gottlob G., Greco G., 2007. On The Complexity of Combinatorial Auctions: Structured Item Graphs and Hypertree Decompositions. In: *Electronic Commerce archive: Proceedings of the 8th ACM conference on Electronic commerce* San Diego, California, USA Pages: 152 - 161

Cereci İ., 2009. CAWP: A Combinatorial Auction Web Platform. MSc Thesis, Atılım University, Ankara, Turkey.

Bakos Y., 1998. The Emerging Role of Electronic Marketplaces on the web. In: Communications of the ACM 41(9), pg 35–42.

Wolfram|Alpha, 2009. eBay as a financial entity. In Site: ”http://www02.wolframalpha.com/input/?i=eBay&a=* C.eBay-_*Financial-“. As of August 2009.

Klemperer P., 2004. A Survey of Auction Theory. In Book: *Auctions: Theory and Practice*. Publisher: Princeton University Press.

Kumar M., Feldman S., 1998. Internet Auctions. In: Proceedings of the 3rd conference on USENIX Workshop on Electronic Commerce - Volume 3 pp.5.

Omote K, 2002. A Study on Electronic Auctions. In: Japan Advanced Institute of Science and Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Brown K.L., Fujishima Y., Shoham Y., 1999. Taming the Computational Complexity of Combinatorial Auctions: Optimal and Approximate Approaches. In: *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence.*

Cramton P., Shoham Y., and Steinberg R., 2006. *Combinatorial Auctions*, The MIT press.

Sandholm T., Suri S., Gilpin A., Levine D., 2001. CABOB A Fast Optimal Algorithm for Combinatorial Auctions. In: *Seventeenth International Joint Conference on Artificial Intelligence (IJCAI).*

Brown K. L., Pearson M., Shoham Y., 2000. Towards a Universal Test Suite for Combinatorial Auction Algorithms. In: *ACM Conference on Electronic Commerce.*

Land A. H., Doig A. G., 1960. An automatic method of solving discrete programming problems. In: Econometrica 28

MySQL, 2005. MySQL Performance Benchmarks: Measuring MySQL's Scalability and Throughput. In: A MySQL® Technical White Paper, March 2005.

Apache, 2009. Apache Performance Notes: In Site: http://httpd.apache.org/docs/1.3/misc/perf-tuning.html As of August 2009.

Brown K. L., Shoham Y., 2009 Empirical Hardness Models: Methodology and a Case Study on Combinatorial Auctions. In: Journal of the ACM, Vol. 56, No. 4, Article 22.

Fang L., Wang Y., 2005. OICAS: An Online Iterative Combinatorial Auction System. In: *IEEE International Conference on Systems, Man and Cybernetics.*

A QUALITY OF CONTEXT DRIVEN APPROACH FOR THE SELECTION OF CONTEXT SERVICES

Elarbi Badidi[1] and Larbi Esmahi[2]

[1]*College of Information Technology, United Arab Emirates University, PO.Box. 17551, Al-Ain, United Arab Emirates*
ebadidi@uaeu.ac.ae
[2]*School for Computing & Information Systems, Athabasca University*
1 University Drive Athabasca, Alberta, T9S 3A3, Canada
larbie@athabascau.ca

Abstract: As wireless networks and mobile devices are becoming ubiquitous, mobile users are increasingly requiring access to application services that can adapt to their context as they move to new locations, for example, in their corporate or partners' networks. The quality-of-context information (QoC) used by these application services is a determinant factor in the adaptation process. As application services typically receive context data from several context services, the selection of suitable context services is of paramount importance in providing mobile users with tailored services. In this paper, we describe our proposed framework for context management and our proposed QoC-based algorithm for the selection of context services. The algorithm takes into account the QoC requirements of application services for each context information to which they subscribe with the Context Broker, on which the framework is relying for context management.

1   INTRODUCTION

As wireless networks and mobile devices are becoming ubiquitous, there is an unprecedented rise in the number of mobile workers who are using a variety of modern handheld devices such as PDAs and SmartPhones to consume online services.

With this proliferation of mobile devices, wireless business applications, i.e. messaging and voice services, and healthcare services are more and more developed and deployed using Web services. Thus, using the Service Oriented Computing paradigm in mobile environments considerably enlarges the range of accessible business applications and enables delivering integrated services across wireless networks.

Mobile workers are increasingly requiring retaining access to services that are similar to their corporate services as they move to new locations in a very straightforward manner without having to configure their working environment explicitly. Therefore, business services should be context-aware to deal with the changing environment of the user. Several definitions of the notion of context have been provided in the literature. According to Dey's definition (Dey (2001)), the amount of information that can be categorized as context information is extremely broad. Location and time are the most widely used context parameters by applications. Wireless devices such as environmental sensors, radio frequency identification (RFID) tags send raw context information such as location, presence and other status information across the network. Specialized services, that we call *context services*, capture, store, analyze and aggregate data to provide high-level context information to application services as needed. Indicators that can essentially be measured and captured include temperature, humidity, pressure, whether the mobile worker is in motion, and many more.

Being context-aware allows application services to make inferences about the current situation of the user using context information obtained from various sources such as surrounding sensors, and GPS receivers, and hence adapt their behavior to the prevailing context of the user. By fusing acquired data and inferring on it, these applications can deduce, for instance, the position or the current user activity such as "user is in a shopping mall" or "user is in a meeting". Such Context-aware application services are being deployed in various industries

such as healthcare, airline transportation, manufacturing, and retail. In the last few years, Context-aware Web services have emerged as a promising technology for building innovative and interoperable context-aware applications (Truong et al. (2009)).

One of the main problems being faced in the area of ubiquitous computing is handling and distributing context efficiently to enhance personalized service delivery to mobile users. The context sources as well as the consumer services of context information are very often physically distributed. For instance, the context sources providing information about the current temperature may be far from the applications that need to adapt their services to the prevailing weather conditions. Furthermore, it is likely that these context sources provide the same context information but with different QoC (Buchholz et al. (2003); Kamran et al. (2008)). Context-awareness raises new challenges like aggregation of context information in a structured format, discovery and selection of suitable context sources.

In this paper, we describe our proposed framework for context management and our proposed QoC-based algorithm for the selection of context services by application services, so that they can adapt their services to the user context. The algorithm takes into account the QoC requirements of the application services for each context information to which they subscribe with the *Context Broker*. The framework is relying on the *Context Broker* for context management. Selection of context-aware application services by users or intermediary entities is not the subject of this work as it has been investigated by several research works (Yu et al. (2009); Kirsch-Pinheiro et al. (2008)). To the best of our knowledge, selection of context services has not been investigated before.

The *Context Broker* implements a *topic-based publish/subscribe* model in which application services subscribe to context information they are interested in, and context services publish their context information. Context Brokers have been used also in the following works (Chen et al. (2003); Bonino da Silva Santos et al. (2007)).

2 BACKGROUND

2.1 Quality-of-Context

Context information is characterized by some properties referred in literature as quality-of-context (QoC) indicators. Buchholz et al. (Buchholz et al.

(2003)) have defined the QoC as: *"Quality of Context (QoC) is any information that describes the quality of information that is used as context information. Thus, QoC refers to information and not to the process nor the hardware component that possibly provide the information."*

Buchholz et al. (2003)) and Kamran et al. (2008) have identified the following QoC indicators: *precision*, *freshness*, *temporal resolution*, *spatial resolution*, and *probability of correctness*.

Precision. This indicator represents the granularity with which context information describes a real world situation.

Freshness. The time that elapses between the determination of context information and its delivery to a requester.

Spatial Resolution. The precision with which the physical area, to which an instance of context information is applicable, is expressed.

Temporal Resolution. The period of time during which a single instance of context information is applicable.

Probability of Correctness. This indicator represents the probability that a piece of context information is correct.

Several competing *context services* may provide the same context information (Buchholz et al. (2003)). Therefore, potential context consumers should be able to select context services on the basis of the QoC they can assure.

2.2 Context Services

A context service typically provides infrastructure support for collection, management, and dissemination of context information concerning a number of subjects. Subjects may be users, objects such as handheld devices and equipments, or the environment of users. The context service acquires context information from various context sources. Sources are usually third parties that collect and provide context information. For example, consider the "temperature" at the current location of the mobile user. This information may be obtained directly from the mobile device of the user. It can also be obtained from a local weather station. Alternatively, it may be obtained from weather TV channels providing weather information nation-wide.

Several research works have investigated the design and the implementation of context services. Shmidt et al. designed and implemented a generic context service with a modular architecture that
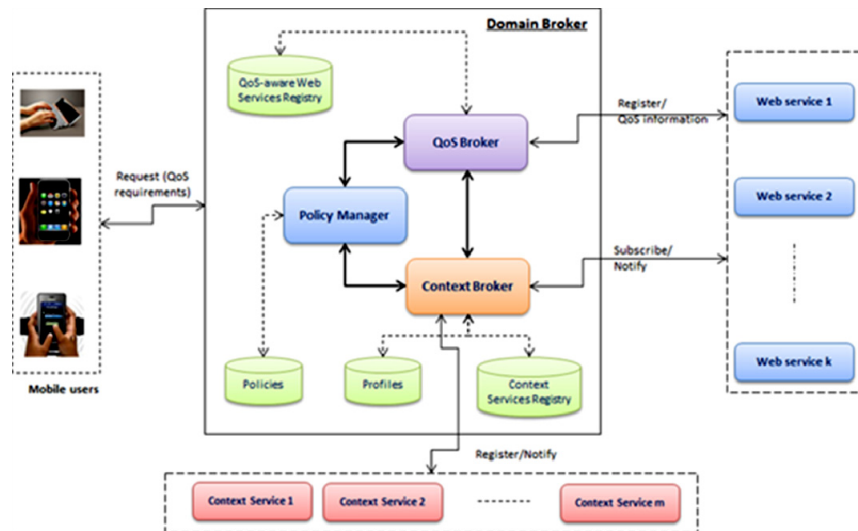
Figure 1: Framework Components.

allows for context collection, discovery and monitoring (Shmidt et al. (2009)). This context service provides a Web service interface that allows its integration in heterogeneous environments. The implementation uses OWL to describe context information and SPARQL to query and monitor context information.

Lei et al. described the design issues and the implementation of a middleware infrastructure for context collection and dissemination (Lei et al. (2002)). They realize this middleware infrastructure as a context service. To allow for wide deployment of the context service, this work has addressed the following issues: Extensibility of the context service architecture by supporting heterogeneous context sources, integrated support for privacy, and quality of context information support. Coronato et al. proposed a semantic context service that relies on semantic Web technologies to support smart offices (Coronato et al. (2006)). It uses ontologies and rules to infer high-level context information, such as lighting and sound level, from low-level raw information acquired from context sources.

3 FRAMEWORK OVERVIEW

As illustrated in Figure 1, our proposed framework is relying on a *Domain Broker*, which mediates between mobile users and application services at the visited site. The *Domain Broker* components are the *QoS Broker* and the *Context Broker*. Another key component of the framework is the *Policy Manager*, which is responsible for managing and maintaining

authentication and authorization policies, as well as polices for monitoring services and their quality of service. The *Domain Broker* components collaborate to deliver personalized context-aware services to mobile users with various devices across interacting sites. The *QoS Broker* is in charge of managing the QoS of Application Web services and submitting the mobile user requests to suitable ones. The *Context Broker* is in charge of managing context information and user profile and preferences. QoS Management operations (QoS specification, monitoring, service level agreement (SLA) negotiation) performed by the *Qos Broker* components are described in our previous work (Badidi et al. (2009)). This paper extends the proposed architecture by considering the context dimension so that applications services are context-aware.

When applications are context-aware, they can adapt their behavior and offer to the user contextually relevant information. Knowledge of the user context allows anticipating the user service and information needs. Context management is achieved in our framework by using a *Context Broker* and by adopting a *topic-based Publish/subscribe* messaging model, a one-to-many pattern of asynchronous message distribution based on registration of interest. Publishers label each message with the name of a topic ("publish") rather than addressing it directly to subscribers. The message system then sends the message to all eligible recipients that expressed their interest in receiving messages on that topic ("subscribe"). The publish/subscribe model is a loosely coupled architecture in which senders often do not need to know who their potential subscribers

are, and the subscribers do not need to know who generates the information.
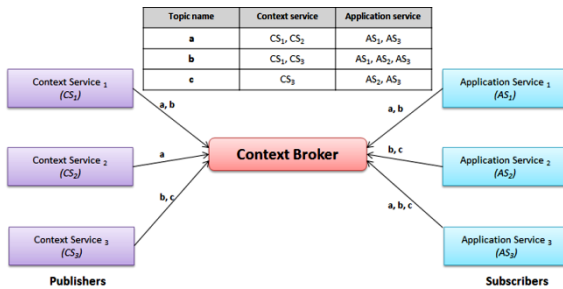
Figure 2: Topic-based publish/subscribe system.

It is increasingly being used in a service oriented architecture context. A Web service disseminates information to a number of other Web services, without the need to have prior knowledge of these other services.

The *Context Broker* is a mediator Web service that decouples context consumers from context services. It is in charge of handling subscriptions of application Web services in which they express their interest to receive context information, and registrations of context services. Once context services are registered with *the Context Broker*, they publish context information and application Web services are then notified by the *Context Broker* about the new context information. Figure 2 illustrates our topic-based publish/subscribe system in which context services are the publishers and the application services are the subscribers. Context information -- such as location, temperature, and user activity -- represents the topics of the system.

4 QoC-BASED CONTEXT SERVICES SELECTION

As we have mentioned earlier, context information may be delivered with different QoC by various context services. Therefore, the *Context Broker* is in charge of selecting appropriate context services to deliver context information to which application services subscribe. Context information may be delivered to the same application service by several context services. Each one may deliver a piece of context information (a topic) that the application service requires to adapt its behavior to the current context.

In the following, we describe the context services selection algorithm for a given application service.

Let $T = \{t_1, t_2, \dots, t_k\}$ be the list of context information (topics) to which an application service has subscribed by showing its interest in receiving such context information.

Let $C = \{CS_1, CS_2, \dots, CS_l\}$ be the list of context services registered with the *Context Broker*. Two context services may provide different context information; each one specializes in offering particular context information. One service, for example, may offer location information while another service may offer only temperature information, and a third one may offer both of them.

These services typically provide context information with different QoC. We assume that QoC indicators are in normalized form with values between 0 and 1. A value of 1 means highest quality and 0 means lowest quality. For example for the *freshness* quality indicator, 1 means that context sources have sensed the information in the last minute, and 0 means that they have sensed it in the last 10 minutes.

When subscribing to context information, an application service specifies the min values of the normalized QoC indicators that it can tolerate. For instance, the application service may subscribe to the *location* information may require a min value of 75% for the *freshness* indicator and 95% for the *probability of correctness* indicator. Let $P = \{P_1, P_2, \dots, P_m\}$ be the list of QoC indicators (parameters) considered in the system.

The minimum quality requirements that the application service tolerates for a given context information (topic) t_j, with $1 \leq j \leq k$, are expressed by the following vector:

$$M_j = \{min_{1,j}, min_{2,j}, \dots, min_{m,j}\}$$

$0 \leq min_{i,j} \leq 1$, with $1 \leq j \leq k$ and m is the cardinality of P.

Therefore, the whole quality-of-context requirements of the application service for all its subscribed topics and all QoC indicators considered in the system can be expressed by the following matrix:

$$Q_a = \begin{array}{c} \\ t_1 \\ t_2 \\ \dots \\ \dots \\ t_k \end{array} \begin{array}{cccc} P_1 & P_2 \dots \dots \dots & & P_m \\ \begin{bmatrix} min_{1,1} & min_{2,1} \cdots & \cdots & min_{m,1} \\ \vdots & & & \vdots \\ min_{1,2} & min_{2,2} \cdots & \cdots & min_{m,2} \\ \vdots & & & \\ min_{1,k} & min_{2,k} \cdots & \cdots & min_{m,k} \end{bmatrix} \end{array}$$

The goal of the selection algorithm is to find for each topic t_j, to which the application service subscribed, a suitable context service from the set C

that can satisfy the minimum quality requirements of the application service.

The QoC offer of a context service CS_r is expressed by the following matrix:

$$Q_r = \begin{array}{c} \\ t_1 \\ t_2 \\ \cdots \\ \cdots \\ t_k \end{array} \begin{array}{c} P_1 \quad P_2 \cdots \cdots \cdots \cdots \quad P_m \\ \begin{bmatrix} q^r_{1,1} & q^r_{2,1} \cdots & \cdots & q^r_{m,1} \\ \vdots & & & \vdots \\ q^r_{1,2} & q^r_{2,2} \cdots & \cdots & q^r_{m,2} \\ \vdots & & & \\ q^r_{1,k} & q^r_{2,k} \cdots & \cdots & q^r_{m,k} \end{bmatrix} \end{array}$$

CS_r is suitable for a topic t_j if the following condition is satified:

$$0 \le min_{i,j} \le q^r_{i,j} \le 1 \text{ for } 1 \le i \le m \text{ and } 1 \le j \le n \quad (1)$$

We define the distance of the QoC offer of CS_r from the application service required QoC for each quality indicator as:

$$d^r_{i,j} = q^r_{i,j} - min_{i,j} \text{ for } 1 \le j \le k \text{ and } 1 \le i \le m$$

Therefore, we can consider the distance matrix $Dist_r$ for the QoC offer CS_r.

$$Dist_r = \begin{array}{c} \\ t_1 \\ t_2 \\ \cdots \\ \cdots \\ t_k \end{array} \begin{array}{c} P_1 \quad P_2 \cdots \cdots \cdots \cdots \quad P_m \\ \begin{bmatrix} d^r_{1,1} & d^r_{2,1} \cdots & \cdots & d^r_{m,1} \\ \vdots & & & \vdots \\ d^r_{1,2} & d^r_{2,2} \cdots & \cdots & d^r_{m,2} \\ \vdots & & & \\ d^r_{1,k} & d^r_{2,k} \cdots & \cdots & d^r_{m,k} \end{bmatrix} \end{array}$$

Using the distance matrix, we can say that the application service requirement can be satisfied for a given topic t_j by the context service CS_r if the corresponding row in the above matrix has values greater than or equal to zero. Therefore, we can discard from that matrix the rows having negative values. We call the resulting matrix $PDist_r$ (all rows have positive values).

The Euclidian distance of CS_r offer from the application service QoC requirement for topic t_j is:

$$d^r_j = \sqrt{\sum_{1 \le i \le m} \left(d^r_{i,j} \right)^2} \quad \text{with } 1 \le j \le k \quad (2)$$

The highest value of d^r_j corresponds to the best QoC offer that can fulfill the QoC requirements of the application service for the topic t_j.

The most suitable context service for topic t_j, that we call here $target_j$, will be the one that maximizes the above Euclidian distance, that is :

$$target_j \leftarrow max_{1 \le r \le l} \left(\sqrt{\sum_{1 \le i \le m} \left(d^r_{i,j} \right)^2} \right) \quad (3)$$

$$target_j \leftarrow max_{1 \le r \le l} \left(\sqrt{\sum_{1 \le i \le m} \left(q^r_{i,j} - min_{i,j} \right)^2} \right) \quad (4)$$

In (4) we have assumed that the application service gives the same weight to all QoC indicators. This is not always the case as the application service may set relative weights for the QoC indicators. The application service may even set weights for each topic to which it subscribed. For example, for the *location* topic, more weight may be given to the *spatial resolution* indicator than to the *probability of correctness* indicator. For the *time of the day* topic, more weight may be given, for example, to the *precision* indicator than to the other QoC indicators. Therefore the weights for a given topic t_j, can be expressed by the following vector:

$$W_j = \left\{ w_{1,j}, w_{2,j}, \dots, w_{m,j} \right\}$$

Where $0 \le w_{i,j} \le 1$ and $\sum_{1 \le i \le m} w_{i,j} = 1$

Given these weights, the most suitable context service for the topic t_j, $target_j$, will be the one that satisfies the condition (1) and which maximizes the sum of quality offers for all QoC indicators:

$$target_j \leftarrow max_{1 \le r \le l} \left(\sqrt{\sum_{1 \le i \le m} \left(w_{ij} \left(q^r_{i,j} - min_{i,j} \right) \right)^2} \right) \quad (5)$$

If no context service satisfies the application service QoC requirements for a given topic, then the Context Broker may ask the application service to lower its QoC expectations.

The key idea of the above QoC-based context service selection algorithm is to find the most suitable context service with regard to the QoC requirements of a given application service for each context information (topic) to which the application service has subscribed.

5 PROOF OF CONCEPT

As a proof of concept, we describe in this section a scenario of how the QoC selection algorithm works. Assume that the topics to which an application service has subscribed are in order of *location*, *time of day*, *activity*, and *temperature*.

Assume that the QoC indicators considered in the system are in order of *freshness*, *probability of correctness*, *temporal resolution*, and *spatial resolution*. The normalized minimum QoC requirements of the application service are given in Table 1.

Assume that four context services have registered with the context broker, CS_1, CS_2, CS_3, and CS_4. In the following we consider only the *location*; the same process applies to the other topics. Table 2 describes The QoC offer of the four Context services for the *location* context information (topic).

Table 1: Minimum QoC requirements of the application service.

	freshness	Probability correctness	Temporal resolution	Spatial resolution
location	0.95	0.75	0.65	0.75
time	0.80	0.85	0.80	0.65
activity	0.65	0.75	0.65	0.50
temperature	0.85	0.90	0.50	0.90

Table 2: QoC offers of four Context service for the *location* topic.

	freshness	Probability correctness	Temporal resolution	Spatial resolution
CS_1	0.97	0.80	0.75	0.85
CS_2	0.80	0.70	0.80	0.65
CS_3	0.98	0.78	0.65	0.80
CS_4	0.95	0.80	0.70	0.80

In this scenario, we consider that the application service assigns the same weight to all QoC indicators. By computing the distance matrix for QoC offers for the *location* topic and the Euclidian distance for each context service, we get the following ranking of the context services from highest offer to lowest:

$$CS_1 (0.1513), CS_4 (0.0866), CS_3 (0.0.0656)$$

The QoC offer of CS_2 does not meet the minimum QoC requirements of the application service.

6 CONCLUSIONS

Context services are increasingly used as an intermediary between context-aware application services and context sources. They provide infrastructure support for collection, management, and dissemination of context information concerning a number of subjects. The adaptation of services to the context of users requires the acquisition of high-quality information from context services. Therefore, the selection of suitable context services is becoming a pressing issue. In this paper, we have presented our proposed framework for context management and our proposed QoC-based algorithm for the selection of context services. The algorithm takes into account the QoC requirements of the application services for each context information to which they subscribe with the *Context Broker*, on which the framework is relying for context management.

REFERENCES

Badidi, E., Esmahi, L., Serhani, M. A., and Ikbal, T., 2009. A QoS-Broker Based Framework for Personalized Web Services Provisioning to Mobile Users, In *Proceedings of the International Conference on Wireless Information Networks and Systems (WINSYS'09)*, Milan, Italy, July 2009, 147-152.

Bonino da Silva Santos, L. O., van Wijnen, R. P., and Vink, P., 2007. A service-oriented middleware for context-aware applications. In *Proc. of the 5th international Workshop on Middleware For Pervasive and Ad-Hoc Computing at Middleware 2007*, pp.37-42.

Buchholz, T., Kpper, A., Schiffers, M., 2003. Quality of context: What it is and why we need it. In *Proc. of the 10th International Workshop of the HP OpenView University association (HPOVUA)*.

Chen, H., Finin, T., and Joshi, A., 2003. An ontology for context-aware pervasive computing environments. *Knowledge Engineering Review* 18: pp.197-207 Cambridge University Press.

Coronato, A., De Pietro, G., and Esposito, M., 2006. A Semantic Context Service for Smart Offices. In *Proc. of the 2006 international Conference on Hybrid information Technology - Volume 02*, pp.391-399.

Dey, A.K., 2001. Understanding and Using Context. *Journal of Pervasive and Ubiquitous Computing*, 5(1), pp. 4–7, 2001.

Kirsch-Pinheiro, M., Vanrompay, Y., and Berbers, Y., 2008. Context-Aware Service Selection Using Graph Matching. In *Proc. of the 2nd Workshop on Non Functional Properties and Service Level Agreements in Service Oriented Computing at ECOWS 2008*.

Lei, H., Sow, D. M., Davis, J. S., Banavar, G., and Ebling, M.R., 2002. The design and applications of a context service. *SIGMOBILE Mob. Comput. Commun. Rev.* 6, 4 (Oct. 2002), pp.45-55.

Schmidt, H., Flerlage, F., Hauck, F.J., 2009. A generic context service for ubiquitous environments. In *Proc. of the IEEE International Conference on Pervasive Computing and Communications (PERCOM 2009)*, pp.1-6.

Sheikh, K., Wegdam, M., and van Sinderen, M., 2008. Quality-of-Context and its use for Protecting Privacy in Context Aware Systems. *Journal of software*, 3(3), pp.83-93, March 2008

Truong, H.L., and Dustdar, S., 2009. A Survey on Context-aware Web Service Systems. *International Journal of Web Information Systems*, 5(1), pp.5-31, Emerald, 2009.

Yu, H.Q., and Reiff-Marganiec, S., 2009. Automated Context-aware Service Selection for Collaborative Systems. In *Proc. van Eck, J. Gordijn, R. Wieringa (eds.): Advanced Information Systems Engineering 2009 (CAiSE'09)*, LNCS 5565, pp. 261-274.

A CONCEPTUAL FRAMEWORK OF SERIOUS GAMES FOR HIGHER EDUCATION

Conceptual Framework of the Game INNOV8 to Train Students in Business Process Modelling

Rebecca Bulander

Pforzheim University of Applied Science, Tiefenbronnerstr 66, 75175 Pforzheim, Germany
rebecca.bulander@hs-pforzheim.de

Abstract: Educational computer games also known as serious games are a new method to teach students. Serious games are considered to be motivation in themselves especially for younger people and seem to be a suitable method to teach students in higher education. The authors of this paper present a conceptual framework of the computer game Innov8 – a business process modelling game to learn the notation BPMN and also how to model business processes. The purpose of the proposed framework is to measure the effectiveness of some aspects of the learning subject business process modelling. Therefore the authors describe the classification of serious games and give a summary about conducted studies in this area of research. After that they describe the framework itself. The paper concludes with a description of further steps of research in the context of a computer course at university.

1 SERIOUS GAMES TO ENRICH EDUCATION

Computer games have become a huge global, cultural phenomenon and also in higher education and training of students. The achievements by using games are enormous, e. g. improved artwork and graphics or immensely impressive physical simulation engines. (Aarseth, 2005). Especially the younger generation is used to this kind of entertainment but also in the context of learning. To some of them, computer games are more important than movies. If computer games don't belong to the violent genres, they can offer opportunities to involve students in special situations and to emphasize specific aspects and interrelations of a given situation while playing. Therefore serious games might be a good opportunity to enrich higher education, to simulate a scenario of the real world and to enhance the didactics for students. Serious games are now taken seriously by scholars and academics (Aarseth, 2005). Serious games can have different positive outcomes. On the one side they are allowing learners to experience situations that are impossible in the real world for reasons of e. g. cost, time or safety (Corti, 2006; Squire, 2003); on the other side serious games can positively influence the learners' development of different skills for example team work (van Eck, 2006). Thinking skills of motivated learners can be described by the following adjectives enthusiastic, motivational, learner-driven, incremental, contextualised, concentrated, interesting or identificational. These attributes are sometimes difficult to generate by a conventional learning session. Serious games themselves can advance these attributes and therefore boost the learning effects (Schwan, 2006). Other advantages of serious games can be to perform real tasks and scenarios in a virtual context, discovery learning through risk taking/failure, immediate feedback and lots of practice. (Clark, 2006)

The development of serious games largely depends on the development of computers, display possibilities, graphical design, interconnectivity and mobility. Based on these advancements and technology development the market of serious games is characterised by high growth rates (Susi, 2007).
In literature we can find several definitions for serious games. Susi et al (Susi, 2007) define them as games that engage the user and contribute to the

achievement of predefined objectives. The adjective serious refers to products used by industries like education, scientific exploration, health care etc. The differences between serious and entertainment games can be summarised, that serious games focus on problem solving of a special task, containing important elements of learning and reflect natural communication (Susi, 2007).

According to the definition above the field to apply serious games is very wide. The focus of the paper is on the programs of higher education where students take specific courses at university in business process management. Besides the general understanding of processes this course aims to educate and trains the students in the process modelling language Business Process Modelling Notation (BPMN). Therefore we selected the software INNOV8 of IBM. This software is specifically for educational purposes and used to train employees and students to the modelling language. Usually BPMN and the phases of business process modelling can be trained by a conventional lecture using slides and also by reading the specific literature to this subject. In this case it depends on the capabilities of the lecturer to impart the knowledge about the notation and also about the whole modelling process. To point out the obstacles of the modelling process in reality e. g. when gathering required information from several departments, especially Innov8, offer a good opportunity to enrich didactics and to demonstrate a specific scenario of reality. Hence the students get the opportunity to have their own experience.

In this paper we will present a conceptual framework of the computer game Innov8. The purpose of the proposed framework is to measure the effecttiveness of some aspects of the learning subject business process modelling. In chapter 2 the authors therefore give a summary about conducted studies in this area of research and describe the classification of serious games. In chapter 3 they describe the framework itself. In chapter 4 the research methodology and design for further steps of research are described. The paper closes with a conclusion.

2 RELATED WORK

In this chapter we will bring up some related work in the field of research of digital game studies. First we provide a classification of digital games. Second we mention some important studies in this field of research. Third we refer to related frameworks.

2.1 Classification of Games

To classify serious games there are two aspects: the commercial sector and the genre of the game. There are diverse industries for which serious games are utilised to educate people (Susi, 2007): Application area, military, government, education, corporate, healthcare and others.

In literature we can find several genres of games. A game genre is the type or category of the game played (Yussof, 2009). A game can be assigned to one or more genres. The following enumeration contains the main common genres (Schwan, 2006):

- Action game: computer game that focuses on physical challenges, including reaction-time and hand-eye coordination e.g. shooting games.
- Adventure game: computer game in which the players assume the role of protagonists in an interactive story that is driven by puzzle-solving and exploration.
- Role-playing game: computer game in which players assume the roles of characters or take control of an avatar in a fictional setting.
- Strategy game: computer game in which players' decisions have a high significance in determining the outcome.
- Puzzle: computer game that emphasizes puzzle solving.
- Simulations: imitation of real world scenarios, state of affairs or process e. g. flight simulation.

For this paper we selected the commercial sector education and in choosing the game Innov8 the game genres role-playing game.

2.2 Learning Principles

Gee (Gee, 2003) summarised 36 learning principles that mean the mechanisms by which players learn in the context of computer games. Some of these learning principles are e. g.: active, critical learning principle, self-knowledge principle or situated meaning principle. Computer games always use several of these learning principles. The learning principles (Clark, 2006) which are mainly used in the serious game INNOV8 are the following eight:

- Active, critical learning principle: Learner-driven, massively participative.
- Semiotic principle: Understanding complex environments and inter-relationships.
- Committed learning principle: Massive motivational commitment through virtual identities and participation in a complete world.
- Amplification of input principle: Massive

amounts of feedback, failure and rewards, often in real time.

- Achievement principle: Intrinsic goals with satisfying achievements and rewards
- Probing principle: The learner has to learn by constantly probing – try things out, test them and try again.
- Situated meaning principle: Performance in meaningful and contextualised – not abstract
- Discovery principle: Narrative is kept to a minimum, forcing the learner to explore and discover.

2.3 Related Digital Game Studies

In literature we find some studies about digital games. However the focus of these studies varies widely. Because of the extended and continuing debate concerning negative effects of (violent) computer games we find a lot of studies dealing with the related matters. Susi et al (Susi, 2007) summarize ten studies about these topics and come to the conclusion that several effects of computer games and also serious games can be measured, but they can't find an evidence of effects related to aggressiveness. The effects that had been measured are: Motor skills, educational and informational, social as well as physiological. As well Susi et al (Susi, 2007) mention that a need for more investtigation in some specific areas concerning the increase but also the decrease in aggressiveness.

Blunt (Blunt, 2008) describes the result of three causal-comparative exploratory studies conducted with the purpose to find out more about the relationship between the use of video games and learning. The result was that the classes using the game had significantly higher means of learning than those classes that did not use the game. Thus these results point out positive effects of the use of serious games.In addition to these results of computer gaming studies further studies can be found. Most of them address special computer games. From this it concludes that the results are difficult to compare. Especially because we know that serious games vary to a great extent in terms of the industry themes they cover, the game genre and the learning principles.

2.4 Related Frameworks of Digital Games

Garries et al (Garries, 2002) describe a generic model about the input-process-outcome framework of games in their paper. The key component of this framework is the game cycle that includes the steps:

user judgement, user behaviour and system feedback. The inputs of the framework are instructional content and game characteristic. The outcome is the learning outcome. The game cycle is an iterative process while the learner is playing the game again and again (see figure 1). One focus of the framework lays on the iterative loop in the game cycle.

The model in figure 1 describes the learning process for serious games in education very well and also fits into modern didactic theories (Schwan, 2006). This model will be considered when we formulate the framework in this paper. An important aspect which is the mission in the model is the motivation of the learner to play a game.

Another very common model was invented by Keller in 1983 (Keller, 1983) and describes a model of motivation design. The model contains four steps for promoting and sustaining motivation in the learning process: Attention, Relevance, Confidence, and Satisfaction (ARCS). The model is very often used in literature and will be characterised in more detail in chapter 3. The ARCS model will also be applied to measure the motivation in learning.
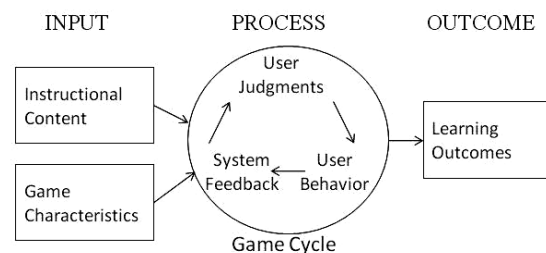
Figure 1: Input-process-outcome framework (Garris, 2002).

3 CONCEPTUAL FRAMEWORK

The following conceptual framework to measure the effects of serious games combines the model of Garries et al (2002) and the ARCS model of Keller (Keller, 1983). The model of Garries contains the steps of a learning process. The ARCS model measures the motivation of the student for learning. Therefore we will give a short overview of the framework and explain single parts of it.

3.1 Overview of the Framework

The framework displayed in figure 2 is an advancement and combination of the two models of Garries et al (Garries 2002) and Keller (Keller, 1983). While Garries et al describes the whole gaming process the

ARCS model reflects on the motivation of the user to play a game and keep on learning while playing.

Another point which was missing is the goals and the objectives to play a serious game. There can be different kinds of goals and objectives which influence the learning process. This aspect is added in figure 2. To measure and interpret the learning effects it is also important to consider what kind of type the player is e. g. frequent player.
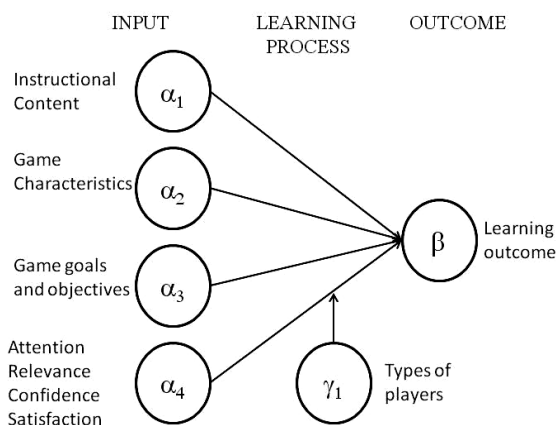
Figure 2: Learning framework for serious games to enrich education.

3.2 Type of the Player

In student education we find a very homogenous group. They are nearly of the same age. But we can still find different types of players. In literature four patterns of types are differentiated (Schwan, 2006):

- Impassioned player: ~ 15 % of all players. These kinds of players are looking for challenges. They have a high frustration tolerance and a high intrinsic motivation.
- Wanna-be-Player: ~ 15 % of all players. These players identify themselves with the impassioned players and want to be like them. They also show a much lower frustration tolerance than the impassioned players do.
- Fun player: ~25 % of all players. These players consider playing of games as one alternative as a recreational activity.
- Occasional player: ~ 45 % of all players. These kinds of players only play a game sometimes, most of the time only as amusement.

The affiliation to one of these different types of players can have an influence on the effectiveness of the outcome of a serious game. Therefore this is also mentioned in the framework.

3.3 ARCS Model

The ARCS model consists of the four elements for promoting and sustaining motivation in the learning process. These four aspects will be explained in the next sections.

3.3.1 Attention

There are many simple ways to win the attention of a learner, but the difficulty lies in sustaining attention. The attention to play a game can be gained in two ways (Keller, 1987):

- Perceptual arousal: to surprise the user and gain his interest.
- Inquiry arousal: to stimulate curiosity by posing challenging questions or problems to be solved.

Therefore different methods for grabbing the learners' attention can be applied like active participation, variability, humor or inquiry. It is most important to find a balance between boredom and indifference versus hyperactivity and anxiety.

3.3.2 Relevance

To make the relevance of a game for a learner clear e. g. for their future career and keep it present in their awareness even if they may be intrinsically motivated, there are several strategies by Keller. One of them is, to point out the learner's future usefulness of the subject; another strategy would be to show the learner, how the new learning will use and extend their existing skills. (Keller, 1987)

3.3.3 Confidence

According to Keller (Keller, 1987) it is very important that one has the feeling of confidence in the possibility of success regardless of external factors or innate ability for a learner. Therefore it may be necessary to provide performance requirements or evaluation criteria and also to establish feedback loops. Another fact is to develop the learner by letting him achieve rising steps of success in the learning process.

3.3.4 Satisfaction

The last element of the ARCS model contains satisfaction of the learning game. The learner must get some kind of satisfaction after a learning period; this can be the achievement of any objective, any praise or entertainment. The learner should get some motivation that the newly acquired skills can help to

solve their problems in a real setting. Thereby care has to be taken that the learner will not be patronized by over-rewarding easy achievements.

3.4 Instructional Content

The instructional content can be described as the subject matter learners should learn by playing the serious game. The subject can vary to a great extent and depend on the objectives and the use as well as the target group of the serious game. Gilbert et al. (Gilbert, 2008) differentiate four types of content: facts, procedures, concepts and principles.

The instructional content of the software Innov8 about business process modelling uses the two content types procedures by addressing business processes and concepts by focussing on the business process modelling notation.

3.5 Game Characteristics

The title game characteristic summarises both topics game mechanics and game rules, by which the details of a game are defined (Thompson et al, 2007). Under the topic game mechanics all technical, graphical and game steering information is summarised; this also contains the user interface and the help function. The importance of game mechanics shouldn't be underestimated because e. g. the attraction of a game with pore graphical design can lose attraction for learners in comparison to other options of games. Also the use e. g. of the help function must be clear and understandable. Otherwise if he gets stuck a learner will be frustrated about continuing the game.

Games take place away from the real world in a fixed space and time period. While playing games the rules and constraints of ordinary life are temporarily interchanged by a set of game rules (Garris, 2002). Rules must be described very clearly and carefully as well as be easy to understand; otherwise the motivation of learners will decrease. The game mechanics of the software Innov8 can be categorised in an average level; it has deficits in the steering of the avatar and the graphical design. The help function contains all the information needed. The game rules are understandable and clearly formulated.

3.6 Game Goals and Objectives

The game goals and objectives are established by the game's rules. The goals contain the criteria of winning and the victory conditions (Blunt, 2008). Ac-

cording to a research of Locke et al (Locke, 1990) clear, specific and difficult goals motivate learners to enhance their performance and engagement; such defined goals allow the learner to compare their achievements during the game and these can be seen as a crucial trigger for greater attention and motivation.

The objectives in the game Innov8 are clearly defined and specific but they can't be seen as very difficult. The learners can always compare their current achievements with the end achievement and results.

3.7 Learning Outcome

The learning outcome is the new skills a learner gained after playing a serious game. The learning outcome can be coupled with the game achievement in playing the game. The learning outcome can be modified based on the game achievement's feedback.In one scenario of the game Innov8 the learner has to redesign the process of a call centre including the staffing with people of different skill levels. The learner can always check in a simulation, what the result of the new process will be, if he left the process in the modelled state like it currently was. When reaching a specific score range, the result of the remodelling of the process will be accepted. The score of the game will be displayed and the game is over.

4 RESEARCH METHODOLOGY AND DESIGN

At the beginning of this chapter we want to point out some important aspects of the proposed framework. Then we will describe the next steps of research. Therefore we will explain the research methodology and design.

4.1 Important Aspects of the Framework

The framework described contains modern didactic theories of the constructivism in which the learner has a central controlling role in the learning process. Learning happens in the willful actions of the learner during a serious game.

Another advantage is that the framework contains all important structures and aspects of modern learning in one model. The ARCS model is also integrated into this framework. Therefore the model doesn't only focus on single aspects like game characteristics or game features.

The framework is adaptable to different kinds of serious games. Previous to the actual adaption all variables have to be checked.

4.2 Research Methodology and Design

The proposed research framework can be transferred in a model which contains hypotheses between the different variables. For example one variable would be instructional content, another one game characteristics. The relationships between the variables are hypotheses which refer to the influence between the variables. The variable "types of player" influences the relationship between the ARCS model and the learning outcome. The aspect if a learner is used to computer games or not makes a difference for the game results and the learning outcome. This model and the hypotheses then represent a structural equation path model. This model can be tested by the Partial Least Squares (PLS) analysis. The PLS procedure was invented by Hermann Wold. It is a second-generation multivariate technique which has the ability to model latent constructs under conditions of non-normality (Chin 1999).

To collect the required data to prove the proposed research model we have already prepared a questionnaire which contains questions for all variables mentioned above (see also figure 1). This questionnaire will be distributed at a university in a class of students who finished their IT laboratory including the lesson of the serious game Innova8. The students are assigned to the course "business administration and engineering". A pre-test has already been conducted and the latest result has been inserted in the proposed model.

5 CONCLUSIONS

In this paper a conceptual framework of serious games for higher education for the game Inno8 has been proposed and explained. The game Innov8 contains learning sessions about the process how to retrieve information of business processes in a virtual company and also how to model and optimize these processes. The result of the game is an optimized process of a call centre. At the end of the game the students can see their result of optimization.

The proposed model is based on modern didactic theories of learning and has a holistic view of the whole learning process. Therefore it takes care of the input and the outcome of the process but also of the game cycle. Consequently the framework also ad-

dresses the aspects that learning from games can be challenging for multiple reasons.

REFERENCES

Aarseth, E. (2005). Game Studies: What is it Good For?. In: The International Digital Media & Arts Association Journal, No. Vol 1, no 3, pp. 3-7.

Blunt, R. (2008): Does Game-Based Learning Work? Results from three recent studies. Defence Gaming – Swedish initiative for linking commercial games with defence. (http://www.defencegaming.org/).

Chin, W. W.; Newsted, P. R., 1999: Structural equation modeling analysis with small samples using least squares. In: Statistical strategies for small sample research. Hoyle, R., Thousand Oaks et al. pp. 307-342.

Clark, D. (2006): Games and e-learning. Caspian Learning Ltd, (www.caspianlearning.co.uk).

Corti, K. (2006) Games-based Learning; a serious business application. PIXE Learning Limited. (www.pixelearning.com).

Garris, R., Ahlers, R., Driskell, J. E. (2002): Games, motivation, and learning: A research and practice model. In: Simulation and Gaming, Vol. 33, No. 4, December 2002, pp. 441-467.

Gee, J. P. (2003): What Video Games Have to Teach Us About Learning and Literacy, Palgrave Macmillan.

Gilbert, L., Gale, V. (2008): Principles of E-Learning Systems Engineering. Oxford: Chandos Publishing.

Keller, J. (1983): Motivational design of instruction. In C. M. Reigeluth (Ed.), Instructional-design theories and models: An overview of their current status Hillsdale, NJ: Lawrence Erlbaum. pp. 383-434.

Keller, J. (1987): Development and use of the ARCS model of motivational design. Journal of Instructional Development. Vol. 10 No. 3, pp. 2-10.

Locke, E. A., Latham, G. P. (1990): A theory of goal setting and task performance. Englewood, Cliffs, NJ, Prentice Hill.

Schwan, S. (2006): Game Based Learning (www.e-teaching.org).

Susi, T., Johannesson, M., Backlund, P. (2007): Serious Games. University of Skövde, Sweden, Technical Report HS-IKI-TR-07-001.

Squire, K.. Jenkins, H. (2003) Harnessing the power of games in education. Insight, Vol. 3, No. 1, pp. 5-33.

Thompson, J., Berbank-Green, B., Cusworth, N. (2007): The computer game design course. Thames and Hudson, London.

van Eck, R. (2006) Digital game-based learning. EDUCAUSEreview, March/April, pp. 16-30.

Yusoff, A., Crowder, R. M., Gilbert, L., Wills, G. B. (2009): A Conceptual Framework for Serious Games. Ninth IEEE International Conference on Advanced Learning Technologies – ICALT 2009: pp. 21-23.

A MEASUREMENT INSTRUMENT FOR INDIVIDUAL INFORMATION COMPETENCY IN AN ENTERPRISE INFORMATION ENVIRONMENT

Chui Young Yoon and Keon Myung Lee

College of Electrical & Computer Engineering, Chungbuk National University, 410 Sungbong-Ro (Gaeshin-dong)
Heungduk-gu, Cheongju city, Chungbuk, 361-763, South Korea
{yoon0109, kmlee}@chungbuk.ac.kr

Abstract: An instrument that can efficiently measure individual information competency is presented to develop and manage the information application ability of individual working in an enterprise information environment. The measurement items are extracted from the major components of a competency. By factor analysis and reliability analysis, a 14-item instrument is proposed to entirely measure individual information capability. The tool's application and utilization are discussed through a case study and the presentation of its results.

1 INTRODUCTION

Today, information technology (IT) progression makes data resources and information systems (IS) become firm's critical strategic resources. (Wu et al., 2005). The efficient utilization of IT in a firm is a critical factor to effectively improve its task performance and competitiveness in an information environment. It is important for human resources working in an enterprise information environment to have the capability to effectively execute the given tasks by applying their information systems to their business. (Mathis and Jackson, 2000; O'Leary, 2002). An individual who directly executes his or her business needs the ability to efficiently perform individual tasks by applying IT and information systems to his or her business in an enterprise information environment.

Hence, this study presents an instrument to measure the individual information competency, which focuses on the entire information capability that an individual can efficiently use information for his or her tasks in an enterprise information environment.

2 THEORETICAL RESEARCH

In previous literature, an end-user was defined as a person who directly interacts with his or her computer. (McHaney et al., 2002; Rondeau et al., 2006; Wu et al., 2007). Based on these studies, we can define an individual as a person who directly interacts with his or her information systems.

In previous literature, competency was defined as effective application of available knowledge, skills, attitudes, and values in complex situations. (Govindarajulu and Reithel, 1998; Bassellier et al., 2001; Tanner, 2001). Namely, the major components of a competency are knowledge, skills, concepts, and development. Individual competency is used to deal with the competence of a person, the collective competency is used to deal with the competence emerging from a group of persons, and global competency is used to describe the organizational ability of an enterprise. (Boucher et al., 2007).

By summarizing prior researches, an individual information competency (IIC) can be defined as the total capability that an individual can efficiently apply information knowledge, skills, attitudes, and values to his or her tasks to execute the given tasks in an enterprise information environment. In other words, IIC is defined as the total capability that an individual directly interacts with his or her information systems to efficiently perform the given tasks through using an organizational data and solutions on information systems. IIC is the entire information capability that an individual can effectively do his or her tasks on an enterprise information system. Based on the definitions and components of IIC, this study generated the 27

measurement items that can gauge IIC in an enterprise information environment. (Govindarajulu and Reithel, 1998; Bassellier et al., 2001; Tanner, 2001; Boucher et al., 2007; Torkzadeh and Lee, 2003; McCoy, 2001).

3 METHODS

Previous literature proposed methods to verify the validity and reliability of the model construct. Most studies presented two methods for a model construct validation: (1) correlations between total scores and item scores, and (2) factor analysis. (Brancheau and Brown, 2002; McClelland, 1973; Boyatiz, 1982; Jacobs, 2002). Torkzadeh and Doll (1999) and Torkzadeh and Lee (2003) used correlation analysis to verify the validity of the model construct. Etezadi-Amoli and Farhoomand (1996), and McHancy et al. (2002) utilized factor analysis to verify the validity of the model construct. We verify the validity and reliability of the instrument construct and extracted adequate measurement items by factor analysis and reliability analysis. The ratio of sample size to number of measurement items (11:1) was above the minimum (10:1) ratio suggested for factor analysis by previous literature. (Torkzadeh and Lee, 2003; Rodriguez et al., 2002). The measurement questionnaire used a five-point Likert-type scale; where, 1: not at all; 2: a little; 3: moderately; 4: much; 5: a great deal. The questionnaire is composed of two response domains: one is answer to general data of respondents, such as degree, age, gender, major field, industry and business department, business position level and years of job experience; the other is response to the measurement items.

3.1 Sample Characteristics

A sample of 243 usable responses was obtained from a variety of industries and business departments, and from management levels. The respondents in terms of business departments were identified as strategy planning (21.1%), development and maintenance (26.8%), business application (38.4%), and administration support (13.7%). The respondents identified themselves as top manager (3.7%), middle manager (44.7%), and worker (51.6%). The respondent had on average of 8.9 years of experience (S.D. =1.118) in their field, their average age was 32.9 years old (S.D. =6.473), and their gender, male (79.8%) and female (20.2%).

3.2 Analysis and Discussion

Items were excluded when their correlation with the collected item-total was < 0.5 or when their correlation with the criterion scales was < 0.6. (Torkzadeh and Lee, 2003; Rifkin et al., 1999; McCoy, 2001). The correlations with the corrected item-total and the criterion item were significant at p ≤ 0.01 and similar to those used by others in previous researches. (Torkzadeh and Lee, 2003; Rifkin et al., 1999; McCoy, 2001). After these analyses, the first 27 measurement items were reduced to 14 items, with 13 items were deleted. The elimination was sufficiently considered to ensure that the retained items were adequate measures of IIC. The validity and reliability of the instrument were verified by factor analysis and reliability analysis. They were used to identify the underlying factors or components that comprise the IIC construct. These deletions resulted in a 14-item scale for measuring IIC. Each of the 14 items had a factor loading > 0.637. The reliability coefficients (Cronbach's alpha) of four potential factors had values > 0.797, above the threshold recommended for exploratory research. (Rodriguez, 2002). The descriptions and loadings for the 14 items are presented in Table 1 and Table 2.

Table 1: Factor loadings obtained from factor analysis.

Variable	Factor Loadings			
	Factor 1	Factor 2	Factor 3	Factor 4
V04	0.794			
V06	0.762			
V03	0.637			
V10		0.881		
V08		0.765		
V13		0.741		
V11		0.699		
V18			0.872	
V12			0.811	
V16			0.794	
V19			0.732	
V17				0.779
V23				0.756
V26				0.659

To examine the reliability and validity of the measures, we calculated the corrected item-total correlations between each variable and its corresponding factor. These correlations along with alpha coefficients of each factor are presented in Table 2. This also shows the alpha coefficients for the measurement of factors if a measure was deleted from the scale. These coefficients indicate the relative contribution of a measure to the construction of a scale for measuring a particular factor. They are all in the acceptable range. Most corrected item-total

correlations were greater than 0.600, showing that the individual measures are good indicators of their corresponding factors.

Table 2: Corrected item-total correlations and the coefficient alphas of 14-measurement items.

Variable	Corrected item-total correlation	Alpha if item deleted
V04	0.681	0.735
V06	0.721	0.791
V03	0.670	0.653

Coefficient alpha for the above 3 items as a composite measure of Factor =0.823

V10	0.812	0.845
V08	0.744	0.812
V13	0.781	0.749
V11	0.789	0.736

Coefficient alpha for the above 4 items as a composite measure of Factor =0.903

V18	0.841	0.817
V12	0.785	0.778
V16	0.821	0.849
V19	0.679	0.712

Coefficient alpha for the above 4 items as a composite measure of Factor =0.813

V17	0.735	0.802
V23	0.819	0.721
V26	0.648	0.612

Coefficient alpha for the above 3 items as a composite measure of Factor =0.797

Hence, the measurement items, with a validity and reliability, were extracted by carrying two analyses as shown in Table 1 and Table 2.

4 MEASUREMENT INSTRUMENT

These analyses classified the extracted items as 4 factor groups. These factor groups indicate the potential major factors to measure the IIC. With investigating the measurement items of each factor, we generated the 4 potential factors as follows: factor 1: information concepts; factor 2: information knowledge; factor 3: information utilization; and factor 4: information development. The 4 potential factors are considered the major factors of the instrument construct. Figure 1 shows the structure of the measurement instrument based on the 4 potential factors and 14 items. Each factor has three or four measurement items as shown in Figure 1.
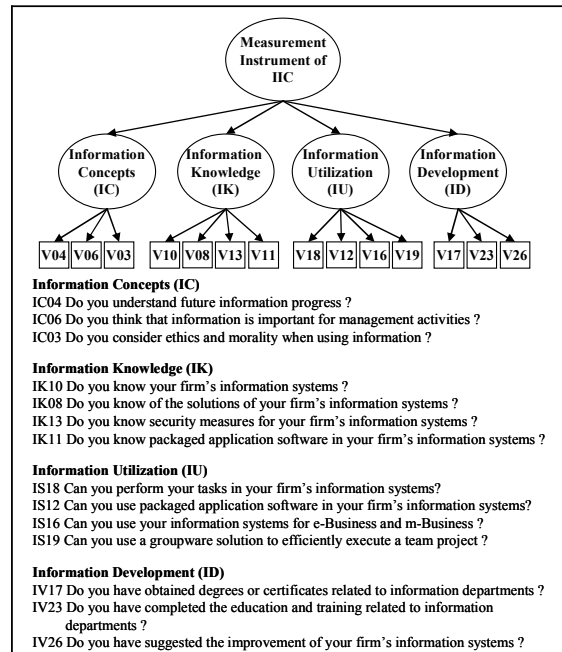
Figure 1: Structure of measurement instrument.

4.1 Measurement Factors and Items

This instrument has 4 major factors to measure IIC in an enterprise information environment.

Information concepts mean state of mind, feelings, and belief related to IT. It includes the measurement items that can identify individual attitude on the future IT progress, IT importance for a firm, and ethic and morality in using information on an enterprise information system.

Information knowledge indicates complex process of remembering, relating or judging information to efficiently use an information system. Namely, information knowledge represents IT knowledge to effectively perform the given tasks on an enterprise information system. It comprises the items that can gauge IT knowledge related to hardware, software, networks, and database for a firm information system, knowledge of packaged application software related to ERP, SCM, and CRM, knowledge related to e-Business (B2E, B2C, and B2B), and knowledge related to security measures in a firm's information system.

Information utilization is the ability that utilizes information to perform specific mental or physical tasks, and includes mental or cognitive skills. Information skills mean the ability that an individual utilize IT knowledge, solutions, and information systems to his or her tasks. It contains the skills as follows: utilization of network and server; use of

packaged application software, such as ERP, SCM, and CRM; use of the information systems for e-business of the form B to E, B to C, and B to B; and the skills to use the security measures in a firm's information system.

Information development refers to the endeavor to improve knowledge and skills related to information. It provides the potential ability to efficiently improve IIC. It has the items that can measure an individual mind on degrees and certificates, domestic and overseas educations and trainings, and suggestion for the improvement of your information systems. This is the important factor for the extension of information capability in terms of the breadth and depth of IIC.

This instrument is a crucial theoretical construct to measure an individual's total information ability that can efficiently do his or her tasks in an enterprise information environment.

5 MEASUREMENT PROCESS

5.1 Framework of Measurement Process

The measurement process has two main stages, including the measurement stage and presentation stage of the measurement results (Figure 2). The measurement stage examines individuals by a questionnaire based on 4 measurement factors and 14 items. The measurement results are analyzed by extracting the measurement values of each factor and by applying each weight value to the measurement values of each factor.
The presentation stage provides the measurement results based on each factor. The results are explained by each measurement index extracted from each factor. The interpretation of the results presents the present states and problems of the IIC, and the directions and methods to efficiently improve the IIC based on the extracted measurement indices.

5.2 Measurement Method

We used the weight values for each measurement factor in order to develop an efficient instrument considered the relative importance of each factor in measuring IIC. The weight values, Table 3, were extracted from the analysis results of the questionnaire survey (AHP) for about 30 experts working in our IT research center.
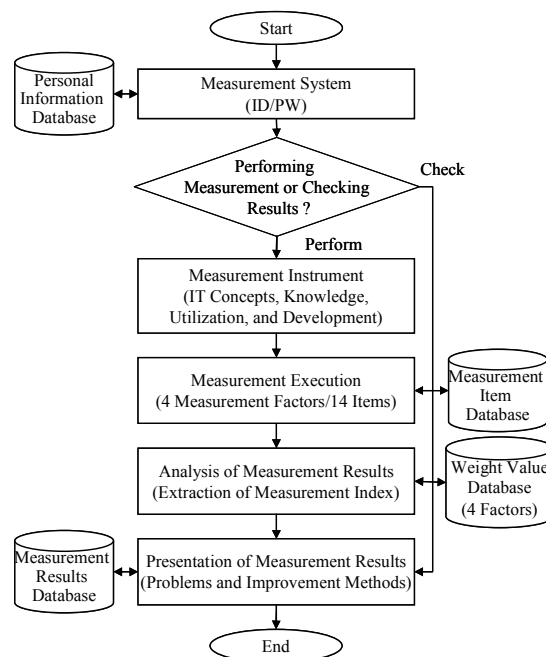
Figure 2: Measurement process.

The measurement method first calculates the measurement values of each factor based on the analysis results that an individual is tested by the measurement items of each factor. It figures out the measurement indices of each factor by multiplying each weight value by the measurement value of each factor.

Table 3: Weight values for each measurement factor.

Measurement Factor	Weight Value
Information Concepts	0.20
Information Knowledge	0.25
Information Utilization	0.33
Information Development	0.22

The measurement index (MI) means the value extracted by multiplying the weight value by the measurement value. The sum of the measurement indices of each factor becomes the individual entire MI.

Hence, the measurement index (MI) of each factor can be presented as Equation (1).

$$MI_{MFi} = MV_{MFi} \times WV_{MFi} \qquad (1)$$

Where, MI_{MFi} : Measurement index (MI) of the i th Measurement Factor

MV_{MFi} : Measurement Value (MV) of the i th Measurement Factor

WV $_{MFi}$: Weight Value (WV) of the i th Measurement Factor

Here, the sum of the weight values of each factor is 1.00 and i = 1, 2, 3 and 4 indicate four measurement factors.

Therefore, the total MI can be defined as Equation (2) by Equation (1):

$$Total\ MI = \sum_{i=1}^{4} MI\ _{MFi} \qquad (2)$$

Here, the i = 1, 2, 3 and 4 mean the four measurement factors.

In this way, this instrument presents the measurement results of IIC based on the total MI and measurement indices of each factor. The problems of the IIC are presented by the results. The efficient methods to improve the IIC are also provided by the total MI and measurement indices of each factor.

6 CASE STUDY AND DISCUSSION

This case study applied the developed tool to 137 workers working in "B" enterprise, Republic of Korea.

6.1 Analysis and Discussion: Overall Organization

We presented the measurement results of each business department and overall organization. The total MI of the overall organization was 62.73, and it was quite high. The business application department (BAD) and the administration support department (ASD) were 65.27 and 63.16 as shown in Figure 3.

The measurement results of each business department represented that the MI of the BAD was higher than those of the other departments. This is due to the ability to effectively accomplish their tasks by frequently applying information knowledge and systems to e-Business of the form B to C, B to B and B to E, and the knowledge and abilities to utilize the various solutions, such as ERP, SCM, and CRM to effectively perform their business tasks on an enterprise information system.
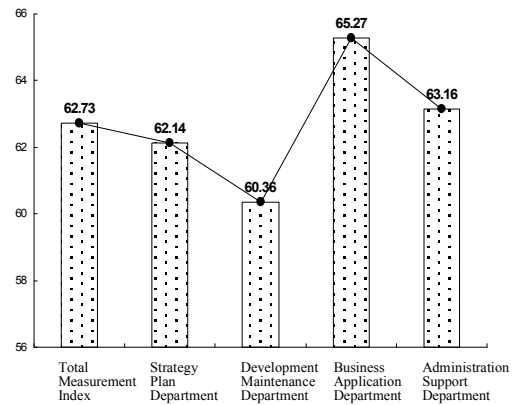
Figure 3: Measurement indices of each business department and overall organization.

Figure 4 presents the measurement indices of each factor for each business department within the organization. The MI of the BAD in all measurement factors was higher than those of the other departments. Especially, the BAD was very high level in information utilization. It indicates that the BAD had the distinguished skills to utilize information solutions and systems to efficiently perform the given tasks. The strategy plan department (SPD) was quite high in the information development. It means that they completed the endeavor to improve knowledge and skills related to information
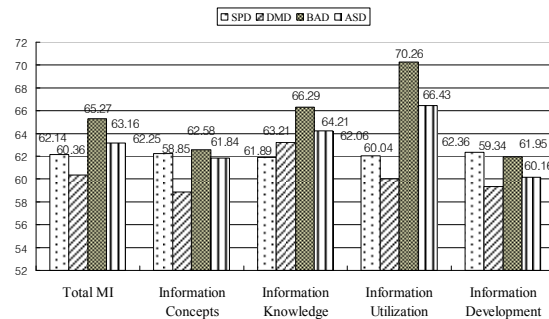
Figure 4: Measurement indices of each factor for each business department.

6.2 Analysis and Discussion: an Individual

The total MI of the individual was 64.89, and it was a little high. Especially, the MI of the information utilization was very high. This means the outstanding skills to utilize the information knowledge, solutions, and systems to his or her tasks in an enterprise information environment. However,

the MI of the information concepts and development were low as indicated in Figure 6.
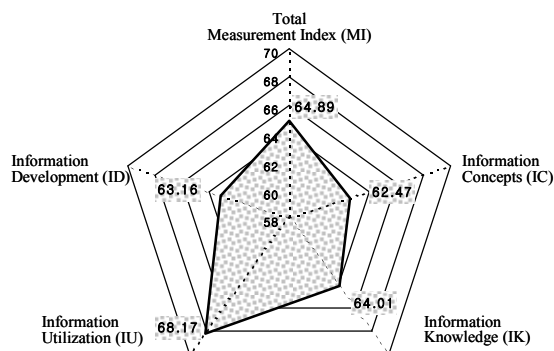
Figure 6: Measurement indices of an individual in the BAD.

Therefore, this individual should make an effort to improve the information concepts and development to effectively raise his or her information competency in general.

7 CONCLUSIONS

We presented an instrument that can efficiently measure an IIC in an enterprise information environment. This instrument includes structure, concrete items, and measurement process and method. This instrument has a nature as a global standard across industries, and business departments and positions.

Therefore, this study provides an instrument that can measure IIC required to efficiently execute an individual's given tasks in an enterprise information environment.

REFERENCES

Bassellier, G., Reich, B. H., Benbasat, I., 2001. Information technology competence of business managers: A definition and research model, *Journal of Management Information Systems*, vol.17, no.4, pp.159-182.

Boucher, X., Bonjour, E., Grabot, B., 2007. Formulisation and use of competencies for industrial performance optimisation: A survey, *Computers in Industry*, vol.58, pp.98-117.

Boyatiz, R. E., 1982. *The Competent Manager: A Model for Effective Performance*, New York: John Wiley & Sons.

Brancheau, C., Brown, V., 2002. The management of end-user computing: Status and Directions, *ACM Computing Surveys*, vol.25, no.4, pp.437-482.

Etezadi-Amoli, J., Farhoomand, A. L., 1996. A Structural Model of End User Computing Satisfaction and User Performance, *Information & Management*, vol.30, pp.65-73.

Jacobs, R. L., 2002. *System Theory Applied to Human Resource Development*, in Theory-to-Practice Monography, Alexandra, ed. D.B. Gardous, VA: ASTD.

Mathis, R. L., Jackson, J. H., 2000. *Human Resource Management*, South-Western College Publishing.

McClelland, D., 1973. Testing for Competence Rather than for Intelligence, *American Psychologist*, vol.28, pp.1-14.

McCoy, R. W., 2001. Computer Competencies for The 21st Century Information Systems Educator, *Information Technology, Learning, and Performance Journal*, vol.19, no.2, pp.101-123.

McHaney, R., Hightower, R., Pearson, J., 2002. A validation of the end-user computing satisfaction instrument in Taiwan, *Information and Management*, vol.39, pp.503-511.

O'Leary, B. S., Lindholm, M. L., Whitford, R. A., Freeman, S. E., 2002. Selecting the Best and Brightest: Leveraging Human Capital, *Human Resource Management*, vol. 41, no. 3, pp.325-340.

Rodriguez, D., Patel, R., Bright, A., Gregory, D., Gowing, M. K., 2002. Developing Competency Models to Promote Integrated Human Resource Practices, *Human Resource Management*, vol. 41, no. 3, pp.309-324.

Rondeau, P. J., Nathan, T. S., Vonderembse, M. A., 2006. How involvement, IS management effectiveness, and end-user computing impact IS performance in manufacturing firms, *Information and Management*, vol.43, pp.93-107.

Tanner, C. A., 2001. Competency-based Education: The New Panacea?, *Journal of Nursing Education*, vol.40, no.9, pp.387-388.

Torkzadeh, G., Doll, W. J., 1999. The Development of a Tool for Measuring the Perceived Impact of Information Technology on Work, *Omega, International Journal of Measurement Science*, vol. 27, pp.327-339.

Torkzadeh, G., Lee, J. W., 2003. Measures of perceived end-user's computing skills, *Information & Management*, vol.40, pp.607-615.

Wu, J. H., Chen, Y. C., Chang, J., 2005. The IS Manager: A Study of Critical Professional Activities and Skills/Knowledge, *Proceedings of the 38th Hawaii International Conference on System Sciences*, pp.1-9.

Wu, J. H., Chen, Y. C., Lin, L. M., 2007. Empirical evaluation of the revised end-user computing acceptance model, *Computers in Human Behavior*, vol.23, pp.162-174.

POSTERS

INTEGRATING MOBILE AGENT AND CONTEXT-AWARE WORKFLOW ANALYSIS FOR M-COMMERCE APPLICATIONS

Jiazao Lin[1], Xining Li[2] and Lian Li[1]

[1]Department of Computer Science & Department of Mathematics, Lanzhou University, Lanzhou, Gansu, P. R. China
[2]School of Computer Science, University of Guelph, Guelph, Ontario, Canada
Hearth84@gmail.com, lil@lzu.edu.cn, xli@cis.uoguelph.ca

Abstract: Mobile commerce (M-commerce) is an attractive research area due to its relative novelty, rapid growth, and great potential in business applications. Unfortunately, there are a number of constraints effecting both performance and usability of mobile devices and network bandwidth. In addition, existing M-commerce applications are lack of fully automated business processes and still require significant manual effort. In this paper we present a general solution of integrating mobile agent and context-aware workflow to implement automated trading task and compose services dynamically in real time to create a highly personalized assistant. Furthermore, the proposed context-aware model derives from a set of ontology of descriptive contextual attribute for knowledge sharing and logical inference. Instead of executing a fully defined process description, the composite workflow can be executed with forward or backward selection of services determined at run time. We have carried out an evaluation experiment. The results show that our proposed solution is feasible and viable.

1 INTRODUCTION

With the development of Internet and its related technologies, the most significant change of our daily lives is the way of conducting business. There is no doubt that Electronic commerce (E-commerce) is the most successful model to explore opportunities and expand business into global commercial market. It offers new channels and business models for buyers and sellers to effectively and efficiently trade goods and services over the Internet （Laudon and Traver, 2006）. However; the traditional E-commerce is based on the client/server approach, which requires a stable connection between client and server. Such a requirement makes some restrictions on the spatial and temporal activities. With various types of Internet-enabled mobile devices, such as PDA's, mobile phones, pocket PC's, etc., the mobile Internet is opening the door to numerous new mobile applications and services that will assist mobile users to engage in time-critical, goal-driven tasks (Sadeh, 2002). M-commerce has emerged and attracted a growing number of research efforts recently. The basic idea of M-commerce is to conduct business transaction with mobile devices and telecommunication/wireless networks, either directly or indirectly (Bai and et al., 2005). It can help nomadic users to roam a wide range of services and products over the Internet on an anywhere and anytime basis. Due to the mobility, personality and flexibility, M-commerce is likely to become the main business model in the near future (Ngai and Gunasekaran, 2007).

One issue which must be considered in the design of M-commerce applications is the limitation of mobile environment and interface between users and applications. In the mobile environment, it is impossible to retain a long connection between the client and service providers (Qiang and Hin, 2002) and it is also unreliable to transfer a huge amount of data between client and server. Comparing to applications based on desktop, the mobile handhold devices have some physical constraints, such as small screen size, limited battery capability, limited storage and computing capacity, and low-bandwidth links with a speed which is slow and varies for different periods of time. In addition, the nomadic users need to frequently check trading opportunity (Mihailescu and et al., 2002), as well as carry out fuzzy and complex information exchange and decision-making tasks. Therefore, it leads to the raise of revenue and creates the risk of missing trade opportunities if the trade time is constrained by the

limited availability of physical access to the service (Kowalczyk and et al., 2003).

Furthermore, It is well-known that an M-commerce transaction involves a sequence of activities, such as negotiation, purchasing, shipment, payment and logistic services. These activities demand additional features of automation and optimization. Existing M-commerce applications are lack of fully automated business processes and still require significant manual effort.

In order to solve the problems mentioned above, we propose the concept of integrating mobile agent and context-aware workflow to implement automated trading task, and compose services dynamically in real time to create a highly personalized assistant. A mobile agent is a self-contained executable entity which is capable of autonomously roaming the Internet to access computing and information resources to carry out user specified tasks. Deploying mobile agents in M-commerce could add automatic and intelligent capabilities to conduct a business transaction (Bǎdicǎ and et al., 2005), and offers mobile users the freedom of connection or disconnection/reconnection functionality to reduce network cost and power consumption. Workflow is the process within a system and the rate at which that happens. Through workflow analysis, we could observe and extract how this process takes place and improve it for efficiency and effectiveness. To achieve user desired goals, a management mechanism can be used to aggregate different business steps into workflows, and automatically chain multiple services together by using planning (Chakraborty and et al., 2005). However, existing workflow management systems lack an appreciation of the content of a business process and do not make decisions based on the nature of information being gathered, that is, many decisions are traditionally made in the process description at design time. In our system, we propose to extend workflow systems to integrate and utilize contextual information relevant to nomadic user to enhance a higher level of automation. For an M-commerce application, the contextual information refers to nomadic user profile and preference which plays a crucial role in the simplification of the interaction between human and the virtual digital world. A general sense of context awareness refers to the ability of an application to discover and take advantage of contextual information (Dey and Abowd, 1999). In our work the context information involves the Person Profile, Environment Profile, Current Activity and Context History. The context-awareness and adaptability accommodated into the workflow is called Context-aware workflow, which is defined as the

process of autonomous and adaptive constructing from atomic services to form a specific, complex task with contextual information. The objective of our research is to design a flexible, adaptable mobile infrastructure to accommodate M-commerce applications.

The reminder of this paper is organized as follows. In Section 2, we will discuss some highlight concepts and methodologies involved in our system, such as mobile agent, context model and context-aware workflow. Section 3 will present our proposed solution and the layered system architecture. In Section 4, we show an experiment to testify the feasibility of the proposed architecture. Finally, we conclude our work and indicate future works.

2 HIGHLIGHT CONCEPTS AND METHODOLOGIES

The dramatic evolution of wireless/telecommunication technologies and mobile computing devices has attracted many researchers to migrate their interests from E-commerce to M-commerce. Several M-commerce platforms have been proposed which use either traditional client-server model or agent-based architecture, such as eAuctionHouse (Sandholm and Huai, 2000) and MAE (Mihailescu and et al., 2002). Other M-commerce researches involve mobile advertising, mobile stock trading, mobile marketing, mobile content distribution, etc. On the other hand, in order to explore automation and optimization of business processes, a number of context-aware applications have been developed to support service composition, such as CAWE (Ardissono and et al., 2007) and CACS (Luo and et al., 2006). However, due to the varity of mobile devices, evolving wireless and telecommunication technologies, heterogeneous platforms and existing and emerging business models, there is still a long way to go in terms of developing the user friendly M-commerce applications. Therefore, in our proposed infrastructure, it will incorporate the concepts of mobile agent, context aware computing, web services, workflow, as well as their related models and methodologies. We will present a brief overview of these researches, and discuss their relevance with our system.

2.1 Mobile Agent Technology

Agent technology is a paradigm for structuring, designing and building systems that require complex interactions between autonomous distributed

components. The agent model has been recognized as a highly effective implementation technique in E-commerce or M-commerce (Kotz and Gray, 1999). Especially the mobile agent paradigm has been deployed as a good candidate to overcome the limitations of connectivity, latency and bandwidth of wireless and telecom networks (Hagen and et al., 1998). For example, a nomadic user can dispatch mobile agents from a handhold device to perform an M-commerce application. Once the application has been launched, the user may disconnect from the network. The execution results can either be sent back by mobile agents through SMS/email, or be collected when the user receives a notice and reconnects to the network. In order to ease the access and participation, reduce costs and improve trading efficiencies, support for automation of decision-making, our proposed system adopts the agent technology and deploys two kinds of agents, namely, stationary agent and mobile agent. A stationary agent always resides at its host and be classified as a home agent or a vendor agent. A home agent is responsible to accept requests from a mobile user and dispatch corresponding mobile agents to invoke a trade transaction, whereas a vendor agent acts as the representative of the vendor to keep track of all transactions, inquires, and possible trade negotiation. A mobile agent represents the "runner", who roams the Internet to carry out the assigned task. Obviously, mobile agents must be lightweight in order to swiftly move across the network. Mobile agents are initialized and dispatched by its home agent to migrate to multiple provider sites and communicate with vendor agents to perform their tasks. Having finished their assigned work, mobile agents may move back to the home server to deliver the execution results.

2.2 Context Model

M-commerce applications have a great demand for context awareness, that is, a need to exploit various information in order to adapt application behaviors. Most existing applications focus on location awareness, i.e., to provide personalized services based on the customer's current position in physical space. Generally speaking, context could be any information that is helpful to characterize the situation of an entity, where an entity can be a person, a place, a physical or a computational object (Dey, 2001). In addition, context could be either explicitly indicated by the user or implicitly extracted from other information sources. Certainly, mobile customers want to find the best deal in an M-commerce environment. The best deal can only be obtained by appropriately combining information

gathered from various shopping services in the physical vicinity. To facilitate the development of extensible and interoperable context aware applications and make contextual data usable and sharable by M-commerce applications, it is essential to have a set of principles for specifying any given context from any domain. To achieve this, a set of well-defined, uniform context models and protocols is required. As a formal representation of entities, ideas, and events, along with their properties and relations within a system of categories, ontology allows sharing a common understanding of information and deriving additional information from what is already known (Uschold and Gruninger, 1996).

Derived from above discussion, we consider the context as the entire collection of entities and their properties that can form a meaningful relationship between mobile users and M-commerce applications. As a result, we define the context dimensions which have to be considered in our system. It consists of a set of elements along four axes, namely, Person profile (Name, Sex, Nationality, Birthday, Social Role, Address and Phone, Friend List), Environment profile(Current Time, Date, Longitude, Latitude, Weather, Temperature, Light, Noise), Context history (Visited Webs and Shopping History) and Current activity. It is worth to note that we only listed a portion of context in our system, the content will extended along with the progress of system development.

2.3 Workflow Management

In M-commerce applications, it is desirable that business activities to be completed quickly with high quality and low cost. A well-designed workflow management system can provide potential competitive advantage to manipulate a series of tasks within a business transaction to produce a final result. In recent years, web services are growing and evolving rapidly in M-commerce applications. The development of new service composition by integrating existing services is generating considerable interests in business communities. There are several typical approaches to the web services composition (Milanovic and Malek, 2004). However, these conventional workflow management models and systems do not provide sufficient characteristics such as automatic composition and adaptability verification. It is therefore important to study how to design a workflow management system which can integrate and utilize context information in the analysis process and activity enactment.

For this purpose, we propose a context-aware workflow management system, as an autonomous

and adaptive process which is able to construct a complex service from atomic activities to achieve a specific task with contextual information. The composition service should be executed with the freedom of forward or backward selection and the service selection is determined at run time instead of executing a fully defined process description. Therefore, we introduce an abstract hierarchy whose higher-level elements describe the tasks to be performed in a generic way. The hierarchy consists of two layers, namely, Abstract Activity and Activity Flow Description (AFD).

An Abstract Activity is a high-level description of the capabilities and categorization of an atomic service with similar functionality. The service functionality is specified in terms of its inputs, outputs, preconditions and effects. Each abstract activity has an associated set of context-dependent implementations representing the alternative courses of the action which the service management model should actuate, depending on the context states. Each Abstract Activity is associated with a service type. There are two types of concrete services in our system, namely, E-service and A-service. An E-service represents a common web service which includes the service provider, the purpose of the service and the method of invocation. On the other hand, an A-service indicates an agent-based service which includes the remote agent host, name of the vendor agent, and the method of agent communication. To invoke an A-service, a mobile agent must migrate to the remote agent server and communicate with the named vendor agent. A vendor agent is a persistent agent acting as the representative of the vendor. It provides services interactively with the mobile agent to implement authentication, query, negotiation, etc.

An AFD indicates the order of a collection of activities without the details of execution. It may contain one or more abstract activities. An abstract activity is replaced by one of the concrete implementations at run time. An AFD can also be predefined, to customize the services requested by the nomadic users. It is the responsibility of workflow management system to determine the actual implementation at execution time based on the context information. Therefore, it is necessary to deploy a conceptual level specification language to describe process product, service and information flows, including the tasks, the dependencies between the tasks and the required roles. In our system, we will employ XML to describe a logical business transaction. The primary reason is the flexibility offered by XML in terms of structured multi-object documents, compact message construction as well as

the wide acceptance of XML as the communication standard for wireless-based applications.

2.4 Service Discovery

Service discovery has been widely studied to allow automatic detection of services, especially web services offered by various service providers. A web service can be invoked by using a specified protocol, such as Simple Object Access Protocol (SOAP), and has an interface described by Web Service Description Language (WSDL) and its related information is published to the Universal Description, Discovery and Integration (UDDI). A service discovery protocol mainly involves dynamically discover and select the best currently available service that fits the need of a specific requirement from user. At this stage, we do not consider the selection rules and selection policies, instead, we focus our attention on a novel approach to discover and select service on the basis of Abstract Activity and contextual information. To get the benefits of the web service standardization and to avoid the redesign of another lookup service we adopt UDDI registry with some extension. Concretely speaking, a new tag <AgentService> is added to a UDDI service registry if the service is an A-Service, that is, the service is implemented by agent-agent interaction.

In this stage, we will illustrate a typical control flow of service discovery and service selection during execution of an M-commerce transaction. To begin with, a lightweight mobile agent is dispatched by the mobile device to deliver a XML-based AFD message to the Home server. Having received the AFD message, the Home Agent will transfer the message to the workflow management system for further processing. The Service Discovery and Selection module extracts atomic activities from the AFD message and uses system defined selection rules and contextual information to search available services from the UDDI server. Based upon the search results, the workflow management system will hook up a concrete implementation to each atomic activity, that is, to invoke a web service if the result returned by discovery module is of type E-service, or to create a mobile agent if the result is of type A-service. An E-service will be invoked by the standard SOAP protocol, while an A-service is carried out by a mobile agent who migrates to one or more remote hosts and communicates with various vendor agents to obtain required services.

3 OVERVIEW OF PROPOSED SYSTEM ARCHITECTURE

To support our proposed approach, we are now in the stage of developing the workflow management prototype system. The main characteristics of the prototype design are adaptability and flexibility. Adaptability is achieved by the context awareness mechanism embedded in the system. Flexibility is obtained by the layered architecture, as presented in Figure 1, where each layer will be wrapped by web service interface. The top layer consists of two function modules which reside in the mobile device.

Portal: This module constitutes the interface of the mobile user and the handhold device. It mainly encompasses different GUI facilities where the user can customize tasks by setting preferences, permission profile and personal information. There are three major functions in this module: Service Information model aggregates a series of abstract activities with similar functions in M-commerce applications. Activity Design GUI is a tool for the mobile user to customize an M-commerce application by setting requirements, preferences, and permission profile. User Information Panel offers functionality to manage user's personal information and contextual information.

Abstract and Connection: This module serves as the communication and operation bridge between the mobile user and the home server and includes two major functions. Abstract Composition Engine is responsible to extract abstract activities from user's specification and transforms the abstract activities to an AFD message. Transform Model Engine is used to transform concrete services into abstractions which are cached in Abstract Activity Centre.

The middle layer is the core of our system, that is, the Context-aware Workflow Management, which involves three main modules:

Service Composition Engine: This module provides the function of translating the well-defined abstract Description Level flow into a concrete workflow in which the required resources will be bound. It manages, controls the workflow, and designates the appropriate service to accomplish the task. Workflow Translation function interprets and decomposes the AFD message came from the top layer into a workflow. Logic Control Repository specifies the business logic of the workflow. XML parser is used to parse the AFD message. Service Management maintains service patterns and templates which are frequently used in M-commerce applications. Service Discovery and Selection communicates with UDDI servers, evaluates application conditions for each candidate service
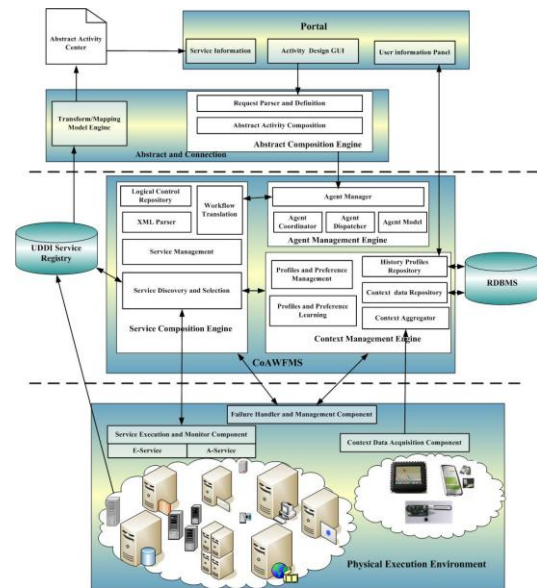
Figure 1: Overview of the System Architecture.

instance and matches appropriate service to an atomic activity.

It will consult Context Manager to bind the context-dependent information to each concrete service implementation.

Agent Management Engine: This module is supported by the IMAGO system [X. Li, 2006].

Context Management Engine: The major function of this module is to gather and process contextual information. History Profile Repository handles every activity performed by the nomadic user across a time span. Context Repository stores all information related to the mobile user and the user's environment. Context Aggregator is used to create new context space based on the existing knowledge and updates. Profiles and Preference Management provides the function of managing the explicit user profile and interest information in a canonical method. Profiles and Preference Learning automatically checks and updates context information through a learning algorithm, such as services and web sites visited frequently.

The bottom layer constitutes the Physical Execution Environment of the home server. In fact, it is a virtual machine specially designed for M-commerce applications.

4 A PROTOTYPE EXPERIMENT

The design methodology and system architecture discussed in previous sections are our ongoing

research project. In order to verify the feasibility of our design, we have conducted some preliminary experiments of system components and functions that are considered to be critical to the viability of our approach. We will use a simple example to demonstrate our experiment. Suppose that a mobile user wants to arrange a trip from Beijing to Shanghai. The trip schedule contains booking a one-way flight ticket, arranging airport pickup, making a hotel and a restaurant reservation. The trip reservation starts at the handhold device of the mobile user. At this stage, we have a simple GUI portal. It should be noted that the prototype at current phase is neither complete nor user friendly. It only serves as a meaningful means with visual representation and guidance to the mobile users for initiating an M-commerce application. In order to cope with context information and context reasoning, we adopt Jena Semantic Web Toolkit to simulate the workflow management module. In the prototype experiment, we defined four classes to simulate and verify the service discovery and activity execution.

Through this experiment, we believe that our proposed approach is capable of adapting existing techniques, such as web service, service discovery, etc., and generating sufficient information for context-aware workflow analysis.

5 CONCLUSIONS

In this paper, we presented the design and architecture which integrates mobile agent technology and context aware workflow to accommodate M-commerce applications. The novelty of our proposal is that it uses an ontological context model to provide personal and environmental contextual information and supports the composition of context-aware services. As a consequence, it not only utilises existing web service and service discovery protocol, but also employs mobile agents to achieve flexible network roaming for interactive services.

Even though we have completed a few critical experiments, the whole research project is still in its very early stage. Our next steps are to complete the workflow management system and to integrate the system with a mobile agent infrastructure. In addition, there are some aspects that should be further investigated. First, we shall study how to model user behaviour though data mining and reasoning, and how to predict the user's actions based on various profiles. Secondly, we will redesign the mobile portal in order to manage the

limited computational resources of handhold devices and provide a user friendly interface. Thirdly, we will investigate the development of M-commerce agents with more intelligent decision-making and learning capabilities in the context of automated business transaction.

ACKNOWLEDGEMENTS

REFERENCES

Ardissono, L. Furnari, R., Goy, A., Petrone, G., Segnan. M, 2007. Context-Aware Workflow Management, In *ICWE'07, Seventh International Conference on Web Engineering,* Como, Italy pp. 47-52.

Bădică, C., Ganzha, M., Paprzycki, M., 2005. Mobile Agents in a Multi-Agent E-Commerce System, In *Proc. Of SYNASC, Timisoara: IEEE Computer Society Press,* pp. 207-214.

Bai, L., Chou, D. C., Yen, D. C., Lin, B., 2005. Mobile commerce: its market analyses, *International Journal of Mobile Communications,* 3(1), 66 - 81

Chakraborty, D., Joshi, A., Finin, T., Yesha, Y., 2005. Service Composition for Mobile Environments, *Journal on Mobile Networking and Applications, Special Issue on Mobile Services,* 10, pp. 435-451.

Dey, A. K. 2001. Understanding and Using Context. Personal and Ubiquitous Computing, *Personal and Ubiquitous Computing Journal,* 5(1).pp. 4-7.

Dey, A. K., Abowd, G. D., 1999. Towards a Better Understanding of Context and Context-Awareness. *Technical Report GIT-GVU-99-22, Georgia Institute of Technology, College of Computing,* Atlanta, Georgia, USA.

Hagen, L., Breugst, M., Magedanz, T., 1998. Impacts of mobile agent technology on mobile communication system evolution, *IEEE Personal Communication,* 5, pp. 56-69.

Kotz，D., Gray, R. S., 1999. Mobile Agents and the Future of the Internet, *ACM Operating Systems Review,* 33(3), pp. 7-13.

Kowalczyk, R., Braun, P., Frankczyk, B., Speck, A., 2003. Deploying Mobile and Intelligent Agents in Interconnected E-marketplaces, *Journal of Integrated Design and Process Science,* 7(3), 109-123.

Laudon, K., Traver, C., 2006. *E-commerce: Business, Technology, Society,* Prentice Hall. Boston, 3rd edition.

Luo, N., Yan, J., Liu, M., Yang, S., 2006. Towards Context-Aware Composition of Web Services, In *Fifth*

International Conference on Grid and Cooperative Computing, Washington, DC, USA, pp. 494–499.

Mihailescu, P., Binder, W., Kendall, E, 2002. MAE: a mobile agent platform for building wireless m-commerce applications, In *8th ECOOP Workshop on Mobile Object Systems: Agent Applications and New Frontiers,* Malaga, Spain.

Milanovic, N., Malek, M. 2004. Current Solutions for Web Service Composition, *IEEE Internet Computing,* 8, pp. 51-59.

Ngai, E. W. T., Gunasekaran, A., 2007. A review for mobile commerce research and applications, *Decision Support Systems*, 43(1) pp.3-15.

Payne, T., Sycara, K., 2002. DAML-S: Web Service Description for the Semantic Web, *Lecture Notes in Computer Science,* 2342(2002), pp. 348-365.

Qiang, W., Hin, H. K. P., 2002. Agent-Based System for Mobile Commerce, In *IPDPS'02, 16th International Parallel and Distributed Processing Symposium*, Florida, pp. 56-60.

Sadeh, N., 2002. *M-commerce: Technologies, Services and Business Models,* John Wiley & Sons. New York, 1st edition.

Sandholm, T. and Huai, Q., 2000. Nomad: Moblie Agent System for an Internet-Based Auction House, *IEEE Internet Computing,* 4(2), pp. 80-86.

Uschold, M., Gruninger, M., 1996. Ontologies : Principles , methods , and applications, *Knowledge Engineering Review,* 11 (2)   pp.93～155.

X. Li, 2006. On the Implementation of IMAGO System, *International Journal of Computer Science and Network Security*, Vol. 6, No. 2A, pp. 107-118

THE IMPACT OF INTERNET ON SOCIAL ANXIETY AND LEARNING ADAPTABILITY FOR YOUNG ADULT INTERNET USERS

Hui-Jen Yang[1], Yun-Long Lay[2], Wen-Yu Tsao[1] and Justin S. Lay[3]

[1]*Department of Information Management, National Chin-Yi University of Technology, Taiping City, Taiwan*
[2]*Department of Electronic Engineering, National Chin-Yi University of Technology, Taiping City, Taiwan*
[3]*Department of Computer Science and Information Engineering, Tunghai University, Taiwan*
{yanghj, yllay, tsaowy}@ncut.edu.tw justin1345@yahoo.com.tw

Abstract: This research is to explore the factors affecting the social anxiety and learning adaptability for young adult Internet users. Results found that achievement expectancy, network self-efficacy, network usage rate and career awareness have significantly affected social anxiety. This result also displayed that achievement expectancy, network usage rate and career awareness have significantly influenced learning adaptability. Social anxiety has a negative effect on learning adaptability. The implications for the academic and educational bureau are also discussed within this paper's closing.

1 INTRODUCTION

Is Internet bringing people more convenient communication or bringing them more anxiety in their daily lives? The network technology has changed the way of people's lives. Individuals probably obsess with net activities, thereby impacting adolescents lifestyle or changes their interpersonal interaction among groups, and then causes social interaction anxiety or learning adaptability problems. However, little evidence has yet come forth in this field. It is worthy of investigation within current college students' social anxiety and learning adaptability.

Several studies have shown that people with people with high social anxiety have poor social capability (Zou, 2007). Faludi (1991) also found people with high social anxiety usually hesitate to communicate with other people Thus, what factors affect college students' social anxiety and learning adaptability are the main issues being examined in this research.

2 THEORETICAL AND HYPOTHESES

2.1 Related Constructs

Clark et al. (1995) argued that fear of negative evaluation and social avoidance is the hallmarks of social anxiety. In this study, social anxiety is defined as the young Internet users' worrying about negative evaluation by other people and fear of interaction with other people. Several studies have shown the factors affecting the social anxiety. The higher the self-focused attention is, the higher the social anxiety is (Kashdan et al., 2004). Expectation and self-awareness have a significant impact on social anxiety, respectively (Bogels et al., 2002). Previous study has also shown that high social anxiety leads to poor social adaptability (Haemmerlie et al., 1988).

Learning adaptability generally includes learning method, learning habit, learning attitude, learning environment and physical adaptability (Li, 1996). Ashforth et al. (2007) presented that the socialization process affects learning adaptability. Career achievement expectation also affects learning adaptability (Kenny et al., 2005). Gong (2003) found that achievement affect learning adaptability. Palthe (2004) noted that self-efficacy affect adaptability.

Bandura (1977) suggested three outcomes of self-efficacy that can predict the changes in people's behavior: choice behavior, effort expenditure and thought patterns and emotional reactions. Thus, in this study Internet self-efficacy focuses on what individuals believe they can accomplish surfing online. Griffin et al. (1998) found that self-efficacy affected the test anxiety. Self-efficacy has a positive effect on achievement and negative impact on test anxiety (Griffin et al., 1998).

Expectancy's theory suggests that the individual will consider the outcomes associated with various levels of performance. Rapee et al. (1997) displayed that social anxiety have an external negative psychological effect on performance. An individual's self-evaluation affects social anxiety. They also presented that task-focused attention would affect social anxiety Rapee et al. (1996) found that self-evaluation had affected social anxiety. Griffin et al. (1998) found that achievement affected text anxiety. Previous evidence also found that achievement affected learning adaptability and social anxiety (McEwan et al., 1999).

Our definition of career awareness is the students' consideration of future job-orientation and its social value. Fadale (1973) developed a Career Awareness Inventory to measure students' career awareness. Career awareness affects life adaptability (Super, 1990). Clark & Wells (1995) wrote that individuals have social anxiety because of their overt self-awareness. Bogels et al. (2002) found social anxiety was affected by self-awareness.

Usage rate is generally classified into light users, middle users and heavy users in the marketing field. In this study, the network usage rate is based on the time that users use the Internet. We classify Internet users into light users, middle users and heavy users. Some evidence has been found to support the impact of network usage rate on anxiety and adaptability. Kaltiala-Heino et al. (2000) found the frequency of involvement in bullying has a significant effect on anxiety.. Ybarra et al. (2007) found that network harassment behavior has a positive influence on psychosocial behavior.

2.2 Hypotheses

Based on the previous empirical research, nine hypotheses are developed and shown below.

H1: network self-efficacy has a negative influence on social anxiety;

H2: network self-efficacy has a positive influence on learning adaptability;

H3: network usage rate has a positive influence on social anxiety;

H4: network usage rate has a negative impact on learning adaptability;

H5: achievement expectation has a negative impact on social anxiety;

H6: achievement expectation has a positive affect on learning adaptability;

H7: career awareness has a positive affect on social anxiety;

H8: career awareness has a positive affect on learning adaptability;

H9: social anxiety has a negative influence on learning adaptability.

2.3 Instrument Development

Responses are provided using a 7-point Likert scale for all constructs, rated from 1 (not at all) to 7 (extremely). The scales of social anxiety, learning adaptability, career awareness, Internet self-efficacy, achievement expectation and network usage rate is adapted and modify from Mattick & Clarke's (1998), Li's (1996), Fadale's (1973), Yang et al.'s (2008), and Ybarra et al.'s (2007) usage scale.

3 RESULTS

3.1 Characteristics of Samples

The mean Internet experience of the young adult respondents was 6.73 years (N=516). The distribution of hours spent on the Internet per week (19.5% under five hours; 22.1% six to ten; 10.6% eleven to fifteen; others are over sixteen hours) and length of Internet usage (continuous use is 30.6% and 43.4% is at least once a day) among respondents indicated that the Internet has become pivotal communication tool for young adult Internet users.

3.2 Reliability and Validity

The alpha value of each construct exceeds the minimum value of 0.7, providing satisfactory reliability as proposed by Nunnally (1978). In this study, the reliability coefficients were 0.88 for achievement expectation, 0.91 for network self-efficacy, 0.78 for network usage rate, and 0.87

for career awareness. All constructs display excellent psychometric properties. Confirmatory factor analysis (CFA) was testing the validity. The CFA is using the EQS software to determine whether the measurement model fitted to our collected data (Bentler, 1995). CFA is conducted using the maximum likelihood method. The measurement model should find χ^2 to be nonsignificant; the goodness-of-fit-index (GFI) should be above 0.90; and root mean square error of approximation (RMSEA) should be below 0.05 (Bentler, 1995). The CFA yields χ^2 /df < 3, p<.0001, NFI=.886, NNFI=.904, IFI=.921, CFI=.920, GFI=.898, and RMSE=0.037 (<0.05). Although the NFI and GFI were slightly below 0.9, these results indicate that the measurement model does fit the data and validate the model proposed by this research.

3.3 Results and Discussions

Structural Equation Modelling (SEM) was used to test hypotheses. The chi-square, RMSEA and a number of goodness-of-fit indices should fit the theoretical requirements (Bentler, 1995). The measures indicate that the model provided a good fit for the data (CFI=.963; NFI=.915; NNFI=.946; GFI=.927; AGFI=.894; RMSEA=.008). Among the nine hypotheses, eight of the nine hypotheses are significant at levels between 0.05, and 0.001, shown in Figure 1. There are four hypotheses involving factors to predict social anxiety. Network self-efficacy and career awareness have a significant impact on social anxiety (β =-.120, P<0.05; β =.524, P<0.001) as the hypotheses predicted. This result indicated that if users have high network self-efficacy, they would have more confidence surfing the Internet. They would thus have no problem in social interaction occasions perceiving the feelings of anxiety. Stronger career awareness would result in higher social anxiety on the part of individuals because young adult will worry about their future and be anxious that they can not survive in society. Clark & Wells (1995) thought individuals have social anxiety because of over self-awareness. Network usage rate and achievement expectation have a direct negative effect on social anxiety (β =-.116, P<0.01; β =-.159, P<0.05) as the hypotheses predicted. These results indicated that young adult Internet users spent too much time on the Internet and may cause social anxiety. The current study also supported the relationship between achievement expectation which conjunction

with past research (Rapee et al., 1997). This result shows that young adult with high achievement expectation will try to solve the problems they face, so they will have low social anxiety.

* P<0.05; ** P<0.01; *** P<0.001
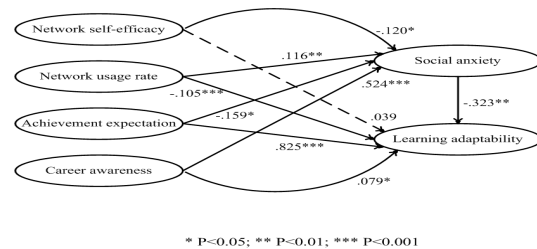
Figure 1 Results of hypotheses.

In turn, five hypotheses involved to predict learning adaptability and all were supported. Network usage rate, achievement expectation, career awareness and social anxiety have a significant effect on learning adaptability (β =-.105, P<0.001; β =.825, P<0.001; β =.079, P<0.05; β =-.323, P<0.01) as hypothesized (H6, H7, H8 & H9). The results indicated that if young adults spend too much time on the Internet, and could cause a learning adaptability problem. Internet users with high achievement expectation more easily adapt to different environments, so they have better learning adaptability. Young adults with higher career awareness would take precautions against the problems they may face planning before they act and simulate the problems they may need to overcome. They will thus have better learning adaptability. Finally, this research also found that social anxiety and learning adaptability have a negative relationship. This means that if young adults have higher social anxiety, they would have less learning adaptability in their daily lives. This result is also consistent with previous research (Zou et al., 2007). However, the influence of network self-efficacy on learning adaptability is not significant (β =.039, P>0.05) unlike as hypothesized (H5) (Griffin et al., 1998; Palthe, 2004). The possible argument is the measurement problem of learning adaptability causing the result.

4 CONCLUSIONS

This study represents an initial effort in mapping some attributes of young adult networking users to social anxiety and learning adaptability. This study

is noteworthy to academic researchers and educators alike. From an academic researcher's perspective, the majority of social anxiety research comes from a medical perspective in studying patients. Little prior research is from the young adult networking users' view to investigate the factors affecting social anxiety. The research results can serve as a reference for advanced study. From the educational perspective, schools have more chance to understand college students' social anxiety and learning adaptability after using Internet in networking society to reduce and prevent the percentage of students from quitting school and enhance their studying adaptablilty.

ACKNOWLEDGEMENTS

REFERENCES

Ashforth, B.E., Sluss, D.M. and Saks, A.M., 2007. Socialization tactics, proactive behavior, and newcomer learning integrating socialization models. *Journal of Vocational Behavior*, 70,447-462.

Bandura, A. 1997. S*ocial learning theory*, Englewood Cliffs, NJ: Prentice-Hall.

Bentler, P. M., 1995. *EQS structural equations program manual*, Encino, CA: Multivariate Software, Inc.

Bogles, S. M. & Lamers, C. T. J., 2002. The causal role of self-awareness in blushing-anxious, socially-anxious and social phobia individuals. Behavior Research and Therapy, 40, 1367-1384.

Clark, D. M. & Wells, A., 1995. A cognitive model of social phobia. In R. G. Heimberg, M. R., Liebowitz, D. A., Hope & F. R. Schneier (Eds.). Social phobia: diagnosis, assessment, and treatment (p.69-98). *NY: Guilford Press*.

Fadale, L. M., 1973. Career awareness inventory. Student Boolet. (ERIC document Reproduction Service No. Ed082062).

Faludi, S. 1991. Backlash: The uncleared war against American women. NY: Crown.

Gong, Y., 2003. Goal orientation and cross-cultural adjustment: an exploratory study. *International Journal of Intercultural Relations*, 27,297-305.

Griffin, M.M. & Griffin, B. W., 1998. An investigation of the effects of reciprocal peer tutoring on achievement, self-efficacy and test anxiety. *Contemporary Educational Psychology*, 23, 298-311.

Haemmerlie, F. M., Montgomery, R.L. & Melchers, J., 1988. Social support, perceptions of attractives, weight

and the CPI in socially anxious males and females. *Journal of Clinical Psychology*, 44(3), 435-441.

Hills, T., Smith, N. D. & Mann, M. F., 1987. Role of efficacy expectations in predicting the decision to use advanced technologies: the case of computers. Journal of Applied Psychology, 72, 307-313.

Hoyt, K. B., 1984. Helping parents understand career education. *Journal of Career Education*, 10(4), 216-224.

Kaltiala-Heino, R., Rimpela, M., Rantanen, P., Rimpela, A. 2000. Bulling at school-indicator of adolescents at risk for mental disorders. *Journal of Adolescent*, 23, 661-674.

Kashdan, F. B. & Robers, J,.E., 2004. Social anxiety's impacts on affect, curiosity and social threat situation. Cognitive Therapy and Research, 28(1), Feb., 119-141.

Kenny, M. E. & Bledsoe, M., 2005. Contributions of the relational context to career adaptability among urban adolescents. *Journal of Vocational Behavior*, 66, 257-272.

Li, Kun-Chung, 1996. Revised scale of learning adaptability, http://teachers.daleweb.org/tools/ z105.htm.

Mattick, R. P. & Clarket, J.C., 1998. Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. Behavior Research and Therapy, 36, 455-470.

McEwan, L. & Goldenberg, D. 1999. Achievement motivation, anxiety and academic success in first year master of nursing students. Nurse Education Today, 19, 419-430.

Nunnally, J. C., 1978. Psychometric theory, 2nd, McGraw-Hill, New York.

Palthe, J., 2004. The relative importance of antecedents to cross-cultural adjustment: implications for managing a global work force. *International Journal of Intercultural Relations*, 28, 37-59.

Rapee, R. M. & Hayman, K., 1996. The effects of video feedback on the self-evaluation of performance in socially anxious subjects. Behavior Research and Therapy, 34, 315-322.

Rapee, R. M. & Heimerg, R., 1997. A cognitive-behavioral model of anxiety in social phobia. Behavior Research and Therapy, 35, 741-756.

Super, D. E., 1990. A life-span, life-space approach to career development. In D. Brown & L. Brooks (Eds.), Career choice and development. Journal of Vocational Behavior, 16, 282-298.

Yang, H. J, Tsao, W.Y., Lay, Y. L., Chen, M., Liou, Y. C., 2008. Prior language experience and language anxiety as predictors for non-native language commercial web site use intention, *International Journal of Human-Computer Studies*, 66, Sept., 678-687.

Ybarra, M. L. & Mitchell, K. J., 2007. Prevalence and frequency of Internet Harassment Instigation: Implications for adolescent health. *Journal of Adolescent Health*, 41, 189-195.

Zou, J. B., Hudson, Jennifer, L., Rapee, R. M., 2007. The effect of attentional focus on social society. Behavior Research and Therapy, 45, 2326-2333.

USE OF WIKI SYSTEMS IN ARCHAEOLOGY
Privacy, Security and Data Protection as Key Problems

Katharina Holzinger, Andreas Holzinger, Christian Safran

Institute for Information Systems and Computer Media, Graz University of Technology, Graz, Austria
katharina.holzinger@edu.uni-graz.at, {a.holzinger, csafran}@tugraz.at

Gabriele Koiner-Erath

Institute of Archaeology, Graz University, Graz, Austria
gabriele.erath@uni-graz.at

Edgar Weippl

Secure Business Research, Vienna University of Technology, Vienna, Austria
EWeippl@sba-research.org

Abstract: Wikis are powerful, collaborative tools and can be used for educational purposes in many ways. The original idea of a Wiki is to make information accessible to all. However, it is very interesting that experiences in the use of Wikis in educational settings showed that security and data protection of wiki contents is definitely an issue. In this paper, we discuss problems and solutions on the basis of use cases from archaeological education. Interestingly, archaeologists are extremely worried about online accessible information due to the serious danger of archaeological looting. "Tomb raiders", i.e. people who excavate artefacts of cultural heritage on the basis of information stored in Geowikis, so called archaeological looters, are not aware of the value of cultural heritage, are interested only in the artefacts and destroying the cultural context, which is of enormous interest for archaeological research. Consequently, the protection of archaeological information is one of the most urgent tasks in the preservation of cultural heritage.

1 INTRODUCTION

Wikis are a Web technology for mass collaborative authoring, offering the asset of free and easy access for end users: everybody can contribute, comment and edit (Gorman, 2005), (Ebner, Kickmeier-Rust & Holzinger, 2008). However, during the discussion after a talk at the conference Archaeology & Computers in Vienna 2009 (Holzinger et al., 2009), it was very interesting that security and data protection of wiki contents is definitely an issue, from the perspective of archaeologists – though neglected by computer scientists. A look into the field of Archaeology makes this obviously evident:

The archaeological heritage is under serious threat from illegal and destructive excavations (archaeological looting) aiming at recovering antiquities for sale on the international market (Brodie & Renfrew, 2005).

2 BACKGROUND

Year by year, thousands of archaeological monuments are destroyed by building projects but also by archaeological looters. In fact, looting is one of the most serious problems in the protection of archaeological heritage. Metal detecting, digging without permission, non reporting of finds and selling illegally excavated objects are forbidden in many countries and – which people are often unaware – are punished severely. A famous example is the looting of several bronze weapons and a 30 cm diameter bronze disk with gold inlays, later known as the sky disk of Nebra with metal detectors on the Mittelberg near Leipzig in Saxony-Anhalt (Germany) in 1999. The disk, a unique object of the 16th century B.C. which is adorned with astral symbols, turned up at the market in 2001, subsequently confiscated by police in 2002, and is

now exhibited in the Prehistoric Museum of Halle/Saale (Meller, 2004).

Looters are fanatical collectors driven by financial motivation or just the ecstasy of finding objects. Looters are mostly interested in metal objects, which can be easily located by metal detectors (Connor & Scott, 1998). By digging up and pulling out the objects, they destroy the archaeological context and thus the possibility to date and interpret the find. Even when these objects are located on the market or in private collections, the damage is irreparable.

Basically, looters do not rely (only) on online information; however, looters are often supported by official find reports that are nowadays widely published in electronic media. These reports offer descriptions and maps of find places and thus make the localization of the monuments very easy.

3 MATERIALS & METHODS

At Graz University of Technology several studies were carried out on how Wiki-systems could be used for different purposes (Ebner, Zechner & Holzinger, 2006), (Ebner & Walder, 2007). In subjects including Civil Engineering, Architecture, Geology, Archaeology, etc. the use of GeoWikis can be of great benefit (Priedhorsky & Terveen, 2008).

Such GeoWiki implementations are focussed on the geographical aspect and are basically implemented as editable maps. Due to specific requirements of the subjects mentioned above, a special GeoWiki, the so called TUGeoWiki has been developed (Safran et al., 2009a), which focuses on the description of the individual locations and the connections between them. The design of the TUGeoWiki server-side and client-side application was based on four central requirements: Geotagging Interface, Mobile and Desktop Interface, Collaborative Environment and Map Mashup (Auinger et al., 2009). The final requirement was to provide the integration of external location-based material, based on the coordinates available for location articles in the TUGeoWiki. This enables the integration of highly detailed material from external sources without the need to provide such material within the TUGeoWiki itself (the mashup concept). The most obvious use for a mashup in the TUGeoWiki's case is the integration of mapping material using an API, such as the Google Maps API. Numerous other map APIs are available and the integration of further material, such as geological data, could also be considered. The final application

consists of two independent systems that fulfil the requirements detailed above. Most of the interaction is performed using a ediawiki, with specially developed extensions within a mobile and/or desktop browser. In order to make this extended functionality available on a mobile device in the field, additional mobile clients for the most common platforms were implemented. On the client side, all of the interaction is done with a browser, which connects to a Mediawiki on the TUGeoWiki server. The original Mediawiki is extended by two special pages. The first offers a list of all geotagged articles in the Wiki. More specifically, this page implements the functionality for the creation of new locations based on geographical coordinates. The second special page supports the upload of images including geographical coordinates. Within JPEG images, additional information can be stored in a so-called EXIF header. Besides metadata regarding a broad range of standard information such as date, time, or camera model, the EXIF header may also contain the geographical coordinates of the image. This information can either be added by a camera, when a photo is taken, or added manually at a later date. The server-side Mediawiki is further extended by a *template*, which is applied to all newly created location articles. The template contains a Google map of the area around its coordinates and it links to the corresponding location on Google Maps and Microsoft Live Maps. Moreover, the template links to another Mediawiki extension called *GeoHack*, which constructs links to various geographical services for the coordinates (Safran, Garcia-Barrios & Ebner, 2009b). Finally, all location articles are added to a common *category*.

4 USE CASES

A recent use case was the Zeğreg Tepe in the territory of ancient Tavium in Asia Minor, a hill where a famous Hellenistic and Roman sanctuary of Zeus Tavianos is supposed to have been located (Strobel & Gerber, 2000), (Koiner, Lohner-Urban & Scherrer, 2010). There was a colossal bronze statue of Zeus, and the sanctuary was known for its asylum, i.e. the protection of people who were persecuted. This sanctuary probably was built on a Hittite temple of Teššub, the Hurritian and Hittite, the weather god who was one of the predecessors of the Greek Zeus. The Zeğreg Tepe is a hill with steep flanks in the south, west and east and a more gentle slope in the north, where the main entrance is to be supposed to be located. The top of the hill is flat

with some pronounced terraces to the south and east. Surveying and air photographs located remains of walls but also rocky structures. Restricting access to these data is mandatory.

5 SOLUTIONS

Restricting access to any information and controlling the use of it are basic security requirements usually implanted in a process referred to as AAA: Authentication, Authorization and Auditing (Weippl, 2005).

Authentication. While there are three basic forms of authentication ("What you are", "What you have", "What you know"), it is mostly passwords, passphrases or similar concepts that are used in Web based applications. The major drawback is that end users tend to select very weak passwords. If combined with a token or certificate ("what you have") authorization becomes more powerful. For Wikis in our specific context, client-side authentication with tokens is feasible. Users have to be given a certificate file that they install on their computer. This certificate is then used to authenticate the client to the Web server.

Authorization. Access Control mechanisms are used to enforce restrictions of authorization. Various security models for access control have been proposed in the last few decades; one can distinguish between discretionary access control (DAC) and mandatory access control (MAC) and role-based access control (RBAC).

DAC controls access to an object on the basis of an individual user's permissions and/or prohibitions. This means that it is at the discretion of an object's owner to decide who has which kind of access. In MAC, the system decides for the user how information should be classified. Changing this categorization to lower the level is commonly prohibited in the Bell LaPadula model (Bell & LaPadula, 1976); the no-write-down property guarantees confidentiality. A role is a collection of operations (on specific objects) needed for an application. Assigning access rights to subjects based on their role is called RBAC. RBAC is commonly used in database systems as it offers two main advantages. First and foremost, RBAC clearly separates "what has to be done" from "who has to do it" in that it assigns users to roles and defines permissions on roles, which in turn takes effect when users activate the corresponding role. Second, RBAC clearly distinguishes an application's object

model from the subject and authorization model (Essmayr, Probst & Weippl, 2004). The object model provides a view on objects to protect (e.g., tables, columns, entities), the subject model highlights which entities are active within a system (e.g., users, processes), and the authorization model describes rules regulating access between subjects and objects and the administration thereof (Weippl, Ibrahim & Winiwarter, 2001). In Wikis with a relatively small user group, role-based access control can be effectively used on two layers. The users and roles can be defined at application level (the Wiki) and, in addition, in the database management system. The Wiki uses, for each user, their database account to connect to the database. While this certainly impacts performance as connection pooling is no longer possible, it clearly increases security. Weaknesses in the Wiki application itself cannot be used to access the database.

Auditing. Analyzing log files to determine who accessed which data is useful to establish whether the Wiki was compromised. Auditing clearly requires reliable authentication and logging mechanisms should be difficult to disable. A good approach is to have two or three levels of logging: (1) Web server logs show the requests made to the Web server. (2) Logs on the application level contain more details on which actions a user performed. However, each application action must have corresponding entry in the Web server log. (3) Logs entries can also be generated inside the database. These entries must again match log entries on the other two levels. If inconsistencies between the log files are detected then an attacker might, for instance, have gained direct access to the database. This can be detected because the database logs would show activities not recorded on the application level. Contemporary Wiki systems only support these concepts rudimentarily, making further development highly necessary.

6 CONCLUSIONS

During the last few years, Wiki systems have become a well-used tool to enable online collaborative work in education. Research has shown the positive effect they can have on the learning process. However, the issues of privacy and security have up to now been neglected in this context, though they are an important factor. A good example is the application of Wiki-based systems in the context of archaeological education. Though especially location-based applications, such as

TUGeoWiki, can enhance the learning process and working in this context, the sensitive nature of archaeological data makes privacy and security a key issue. Authentication, authorization and auditing processes have been shown to be necessary inclusions in the future development of Wiki systems in this context. In this paper, we proposed the implementation of a dual authentication for Wikis to protect information stored in archaeological Wikis. Improving authentication and authorization will enable researchers to share more information without the fear that looters will gain unauthorized access to data and abuse the data. Moreover, precautions can be taken to avoid unintentional publication of hidden data such as location information in images. Unprotected archaeological monuments are highly susceptible to destruction by looters, consequently data protection in this field is highly important.

REFERENCES

Auinger, A., Ebner, M., Nedbal, D. & Holzinger, A. (2009) Mixing Content and Endless Collaboration – MashUps: Towards Future Personal Learning Environments. In: *Lecture Notes in Computer Science (LNCS 5616)*. Berlin, Springer, 14-23.

Bell, D. & LaPadula, L. (1976) Secure Computer System: Unified exposition and multics interpretation. *Technical Report ESD-TR-75-306, 2997 The MITRE Corporation, Bedford (MA)*.

Brodie, N. & Renfrew, C. (2005) Looting and the world's archaeological heritage: The inadequate response. *Annual Review of Anthropology,* 34, 343-361.

Connor, M. & Scott, D. D. (1998) Metal detector use in archaeology: An introduction. *Historical Archaeology,* 32, 4, 76-85.

Ebner, M., Kickmeier-Rust, M. & Holzinger, A. (2008) Utilizing Wiki-Systems in higher education classes: a chance for universal access? *Springer Universal Access in the Information Society,* 7, 4, 199-207.

Ebner, M. & Walder, U. (2007) e-Learning in Civil Engineering - Six Years of Experience at Graz University of Technology. In: Rebolj, D. (Ed.) *24th W78 Conference.* Maribor, Slovenia.

Ebner, M., Zechner, J. & Holzinger, A. (2006) Why is Wikipedia so Successful? Experiences in Establishing the Principles in Higher Education. *6th International Conference on Knowledge Management (I-Know'06).* Graz, Austria.

Essmayr, W., Probst, S. & Weippl, E. R. (2004) Role-based access controls: Status, dissemination, and prospects for generic security mechanisms. *International Journal of Electronic Commerce Research,* 4, 1, 127-156.

Gorman, G. E. (2005) Is the wiki concept really so wonderful? *Online Information Review,* 29, 3, 225-226.

Holzinger, K., Safran, C., Ebner, M., Kappe, F., Koiner, G. & Holzinger, A. (2009). *Geo-Tagging in Archaeology: Practical Experiences with the TUGeoWiki* International Congress Cultural Heritage and New Technologies: Computers and Archaeology, Vienna, 13.

Koiner, G., Lohner-Urban, U. & Scherrer, P. (2010) Die Arbeiten des Grazer Teams in Tavium/Tawinija 2009,Tavium (Büyüknefes, Provinz Yozgat) und seine Region, Bericht über die Kampagnen 2006-2009 (forthcoming). *Istanbuler Mitteilungen,* 60.

Meller, H. (Ed.) (2004) *Der geschmiedete Himmel: die weite Welt im Herzen Europas vor 3600 Jahren. Begleitband zur Ausstellung: Landesmuseum für Vorgeschichte, Halle/Saale (in German),* Stuttgart, Theiss.

Priedhorsky, R. & Terveen, L. (2008) The computational geowiki: what, why, and how. *Proceedings of the ACM 2008 conference on Computer supported cooperative work.* San Diego, CA, USA, ACM.

Safran, C., Ebner, M., Kappe, F. & Holzinger, A. (2009a) m-Learning in the Field: A Mobile Geospatial Wiki as an example for Geo-Tagging in Civil Engineering Education. In: Ebner, M. & Schiefner, M. (Eds.) *Looking Toward the Future of Technology-Enhanced Education: Ubiquitous Learning and the Digital Native.* New York, IGI , 444-454.

Safran, C., Garcia-Barrios, V. M. & Ebner, M. (2009b) The benefits of Geo-Tagging and microblogging in m-Learning: a use case. *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era.* Tampere, Finland, ACM.

Strobel, K. & Gerber, C. (2000) TAVIUM (Büyüknefes, Provinz Yozgat) – Ein regionales Zentrum Anatoliens. Bericht über den Stand der Forschungen nach den ersten drei Kampagnen (1997-1999) (in German) *Istanbuler Mitteilungen,* 50, 215-265.

Weippl, E., Ibrahim, I. K. & Winiwarter, W. (2001). *Content-based Management of Document Access Control.* 14th International Conference on Applications of Prolog, 78-86.

Weippl, E. R. (2005) *Security in E-Learning (Advances in Information Security).* New York, Springer.

THE INFLUENCE OF AJAX ON WEB USABILITY

Andreas Holzinger, Stefan Mayr

Institute of Medical Informatics, Statistics and Documentation (IMI), Research Unit HCI4MED
MedUni Graz & Institute of Information Systems and Computer Media (IICM), TU Graz, Graz, Austria
a.holzinger@tugraz.at

Wolfgang Slany

Institute for Software Technology (IST), Technische Universität Graz, Graz, Austria
wolfgang.slany@tugraz.at

Matjaz Debevc

Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia
matjaz.debevc@uni-mb.si

Abstract: In this paper we discusses some pros and cons of using AJAX for increasing the usability of Web applications. As AJAX allows Web applications look like desktop applications, it can increase the learnability of a Web application. Nevertheless, AJAX can also be the source of end user frustration if the XMLHttpRequest is not supported by the browser, Javascript is not available, or an Internet connection is missing. We also provide some workarounds for server response time gaps, for example by providing visible user feedback messages) and enabling the back button to work properly.

1 INTRODUCTION

The term *user generated content* describes many of the phenomena which became rising stars including MySpace, Facebook, Twitter, or YouTube. All of these so called Web 2.0 sites act only as *interfaces* between publishers and consumers. For both user groups an easy to use interface has to be provided to encourage the publisher to share their content on the one hand and enable an easy access to the uploaded items for the consumers. When the amount of interactions between end users and web pages began to rise, the usability of such websites were in the focus of many companies. Due to the fact that a lot of those services were copied and everyone tried not to lose the contributors to the competitors by improving the quality of the services. With the appearance of the first AJAX approaches such improvements in the field of interactions were possible for the first time. Application programming interfaces and mashups make it difficult to determine between a service, an application or a web page. These facts constitute new challenges for HCI&UE, since every user can be a designer, every menu can have a different behaviour, experience outranks efficiency, and connectivity replaces consistency (Dix & Cowen, 2007).

2 FACTORS OF USABILITY

As stated in ISO-9241-11 usability is "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." This definition makes it impossible to find a universally valid usability recipe: *Usability always depends on external factors (the context)*. For example, a lot of help messages may make an interface more usable for novice users but could seriously disturb expert users in using the interface efficiently.

Thus, all the findings in this paper have to be seen in the *context of the particular usage*. Basically, there are five dimensions of usability: Learnability, Efficiency, Memorability, Error Handling and Satisfaction (Nielsen, 1993), (Shneiderman, 1997), (Holzinger, 2005).

3   AJAX IMPROVING USABILITY

Example 1: The simplicity of HTML ensured that a huge number of people were able to learn and use the language without problems. On the other hand, the simplicity was also its' big disadvantage. Only a small number of tags were available. Such a page normally consist of all design and navigation elements as well as content and links to refer to other pages within the same homepage or to other, third party resources. Even if this never really changed, there is no consistency within different web pages at all because web pages are designed rather than developed. So the position of – for example – navigation elements is up to the creator of the webpage and follows no rules in most cases. While it is mostly agreed that the main navigation elements should be on the left hand site, this is ignored by many sites. Web applications are different due to the combination of hypermedia and an information system and the way to access the data. The way AJAX communicates with the server gave the developers the opportunity to create web applications with the same look and feel as desktop applications for the first time. As user events like pressing a button or dragging a item to the shopping basked can be processed asynchronous in background the application is not interrupted or reloaded in any way anymore (Garrett, 2005).

As Quesenbery (Quesenbery, 2002) stated, the **learnability** of a product can be increased by transferring prior knowledge from similar products to the new one and keeping interaction patterns which the users may have experienced. Mochaui (Houston, 2009) for example is an AJAX framework and enables web designer to create web pages which act similar to the Windows operating system. When entering the page a desktop with clickable icons are presented. A click on those icons opens a window and the requested content will be loaded.

The mental model of using a Windows desktop and opening different software programs to do different jobs is so transferred to a web page in an easy way since most of the functionality of desktop icons and the desktop itself is already implemented in the framework. AJAX is used in the Mochaui framework specially for modifying the DOM model and the style rules defined in the cascade style sheet (CSS) on the fly to enable such interaction like drag and drop and loading content into virtual windows. Those windows are not POP UP windows but only elements of the DOM model and therefore accessible from all other parts of the current loaded page. With this technique even updates to opened windows are possible. It has to be mentioned that all of the features described here in this section would be also possible without Mochaui but this framework combines a bunch of JavaScript methods for easily creating such environments.

Another popular example and one of the first AJAX driven web application with worldwide attention was Google Maps where retrieving a new part of the map is possible by simply dragging the current map in the desired direction. While a non AJAX page would need to reload the entire page, Goole Maps only reloads the graphics for displaying the new section.

The **efficiency** of using the map is so increased by reducing the loading time. Especially when focusing on **memorability** another big advantage of AJAX based web applications becomes visible. As the AJAX fundamental XMLHttprequest is a de facto standard and supported by all major browsers (Firefox, Safari, Opera, Konqueror and Internet Explorer) a AJAX application will work across many platforms (Windows, MacOS, Linux).

Even when some mental models of interacting with the operating systems differ completely the behaviour of the web application can be the same across those platforms. Users which have to do the same task on different platforms – for example when retrieving emails at home, on mobile and at work with 3 different operating systems – the usage of one common web based email client will always offer the same look and feel and the same interaction model. The time effort for a reorientation on the other platform is so minimized or even dropped and increases the memorability especially when one of the platform is used not frequently. The "one application for all platforms" approach could also increase usability in a completely other way: as such an application has only to be developed once, usability tests and removing functional bugs can fulfil the request to minimize **errors**.

Even if only one application has to be developed, it is essential to test it on all target platforms due to the fact that the XMLHttpRequest – Method is not a standard and may be interpreted slightly differently from browser to browser (Asleson & Schutta, 2005).

A similar problem arises when using CSS, where some browsers interpret the CSS rules in a different way.

Example 2. As efficiency can be seen as a relation between the accuracy and completeness of achieving a goal and the resources needed for this achievement (Frøkjær, Hertzum & Hornbæk, 2000), all AJAX gimmicks which reduce the completion

125

time for a certain task can raise the level of efficiency.

A further example are search queries: As search results are always as good as the quality of the queries (Croft & Thompson, 1987) an efficient search system should provide help in finding adequate search terms or phrases. Such an approach is for example Google Suggest, where the system delivers suggestions for expanding the entered search phrase of the user. A drop-down offers the suggestions, which origin from analyzing former search queries. The user simply clicks on one of the entries of the list to finish the search query.

As such methods increase the cognitive load (Bruza, McArthur & Dennis, 2000) such lists are often shorten to the best matching 10 entries; an experiment comparing normal search to a system which helps expanding the query phrases by providing a suggestion list of expanded phrases showed that the relevance of the found pages was higher. Also the time spend on those sites was higher so they matched the expectations more precisely. This technique of providing suggestions is normally done by observing events of an text input field (key down, key up, on-change) to trigger the data retrieval from the server. If a certain amount of characters was typed in, the results set of possible phrase extensions is loaded from the server and displayed in a drop-down window, where a suggestion can be selected. If the user ignores the suggestions and types more words, the suggestions are reloaded and another set is presented to the user.

Another way is to reconfigure and refill elements of a form due to information which was filled in previous fields. During a registration process there is a the request to provide the region where the user lives.

By selecting the appropriate nation, only the regions of that country have to be loaded and placed in the regions drop down. Such lists only consist of alternatives which are necessary for fulfilling the task. It is obvious that finding a term within a shorter set of terms takes less time than going through a huge list of possible answers.

There are two ways of providing such a form reconfiguration: to reduce server load the possible result sets can be loaded along with the page itself. AJAX is only used to tack the trigger events (selection of a nation) and to replace one element of the DOM. This works efficiently when only a few possibilities exists because otherwise all data has to be loaded at once which would increase loading times.

The second approach is to load the data only when the trigger announces a new event.

4 AJAX AS USABILITY KILLER

One of the worst cases when providing a software is when people cannot use the software at all. Unfortunately there are three situations where an AJAX driven application simply does not work.

As the XMLHttpRequest was implemented for the first time in Microsoft Internet Explorer 5.0, Mozilla 1.0 and Safari 1.2 (Asleson & Schutta, 2005) all user with previous versions are not able to use these new websites. Nowadays this group of people is extremely small, because all the supporting versions were published at least 5 years ago. The second smallest group are those users where JavaScript is not available. This can be due to a security policy in companies, when users have scripting turned off or when the browsers don't support scripting (some mobile browsers). According to a statistic of w3schools.com less than 10% of the users have no JavaScript support at their browsers. The third group is probably the biggest, even when the size is not directly measureable: the "offline people" (Nemedi, 2007). AJAX based web applications always need an internet connection to serve the desired data. Even just starting such an application is – with exception of a local cache version – not possible. How many people cannot use the service because they are offline can hardly be predicted but there are several reasons for using a computer in offline mode. Especially in countries where internet access is still a problem, web applications like a online spreadsheet tool are unavailable for those people. When data is reloaded asynchronously and the content of the page is changed without reloading the webpage, the link stated in the browser's address bar does not represent the current state of the application. Depending on the quality level of the implementation a bookmark would only lead to the initial state of the website regardless when the bookmark was set (Kasemvilas & Firpo, 2009). Another issue caused by not having the current state of a website is the wrong behaviour of the back button (Kasemvilas & Firpo, 2009). When for example a user navigates through the pages of a search result by an AJAX request the browser does not recognize that a new page is loaded. Pressing the browsers back button will not lead to one of the previous result pages but to the page that was opened before the entire search. The asynchronous data transmission can also lead into problems when there is a timely gap between the user action and the

corresponding result, especially if the result is displayed on the screen at an unexpected time.

The transmission can even slow down the whole application and confuse the user about the current state when a lot of user interaction takes place which has to be communicated to the server. A common example is evaluating a form by sending every user input to the server and displaying erroneous input. If the location for displaying the error messages is badly chosen, the user may not recognize where the error occurred. Even when there is no error, a time gap between the click and the expected result may lead to confusion and another click on the element so a checkbox is unselected again or an action is executed twice. All of the mentioned problems above may be solved by using AJAX in a well-thought-out way. It is important to use this technology not only by replacing some features of the site with a corresponding AJAX function because that is mostly the reason for site effects as stated above. A workaround for the back-button and the bookmark issue for example is to modify the document location by adding an anchor point (#) to the document.location.href variable. This will lead to an update of the URL in the address bar which can be bookmarked or used as for handling the back-button event. Large web applications inlcuding Google Mail and Facebook use this technique for example to avoid navigation problems. A disadvantage of inserting anchor points is that all major search engines cut off such anchor points and index the wrong content.

5 CONCLUSIONS

AJAX enables developers to design web pages with a complete different look and behaviour than classic pages. Especially the way how users interact with the system helps improving usability by enabling the transfer of already known and learned mental models to web pages. Having only one application for all platforms and being able to remove bugs and usability problems without user notice or user interaction can improve the usability of a site. When a page is reloaded without updating and reformatting the URLs in the address bar, this address may not reflect the current state and point only to the initial state of the web site. Especially bookmarking a specific state of the site is then impossible. When designing an AJAX application, such problems should be taken in consideration already in the design phase. The user should always be informed about the status and – in special cases – that the system is locked because of waiting for data from the server. This avoids overlapping requests to the

server which can fast lead to confusion and unwanted results. For users which are not able to access JavaScript applications due to the lack of the scripting language an alternative should be provided like in Google Mail, where a non JavaScript version is also available. The mail application of Google also handles the problem of users which want to access their emails offline by providing an interface (POP3 / IMAP) for downloading the emails with an offline email reader (email client). As demonstrated, the usage of AJAX may improve the usability, however, the disadvantages should always be kept in mind when designing such applications to avoid a serious step back in usability engineering.

REFERENCES

Asleson, R. & Schutta, N. (2005) *Foundations of AJAX.* Apress.

Bruza, P., McArthur, R. & Dennis, S. (2000). *Interactive Internet search: keyword, directory and query reformulation mechanisms compared.* ACM New York, NY, USA, 280-287.

Croft, W. B. & Thompson, R. H. (1987) A new approach to the design of document retrieval systems. *J. Am. Soc. Inf. Sci.,* 38, 6, 389-404.

Dix, A. & Cowen, L. (2007) HCI 2.0? Usability meets Web 2.0. *Proceedings of HCI 2008.* University of Lancaster, United Kingdom, British Computer Society.

Frøkjær, E., Hertzum, M. & Hornbæk, K. (2000). *Measuring usability: are effectiveness, efficiency, and satisfaction really correlated?* , ACM New York, NY, USA, 345-352.

Garrett, J. J. (2005), Ajax: A new approach to web applications. Technical report. Online available: http://www.adaptivepath.com/publications/essays/archives/000385.php/, last access: 2010-02-22

Holzinger, A. (2005) Usability Engineering for Software Developers. *Communications of the ACM,* 48, 1, 71-74.

Houston, G. (2009), A Web Applications user Interface Library. Online available: http://mochaui.com/demo/demo-virtual-desktop.html, last access:

Kasemvilas, S. & Firpo, D. (2009). *Effects of AJAX Technology on the Usability of Blogs.* Springer.

Nemedi, I. (2007) Ajax Mistakes. *Informatica Economica Journal,* 11, 2.

Nielsen, J. (1993) *Usability engineering.* Morgan Kaufmann.

Quesenbery, W. (2002) Dimensions of usability. *Content and Complexity, Erlbaum.*

Shneiderman, B. (1997) *Designing the User Interface. Strategies for effective Human-Computer Interaction, Third Edition.* Reading (MA), Addison-Wesley.

SMES IN DEVELOPED AND DEVELOPING COUNTRIES TREADING SIMILAR PATHS TOWARDS ETRANSFORMATION

Mahesha Kapurubandara

Sri Lanka Institute of Information Technology, Colombo, Sri Lanka
mahesha.k@sliit.lk

Ana Hol, Athula Ginige

School of Computing and Mathematics, University of Western Sydney, Sydney, Australia
{a.hol, a.ginige}@uws.edu.au

Abstract: The global economy has become increasingly reliant on Information and Communication Technology (ICT) toolsand systems to process and exchange information. The Small and Medium Enterprises (SMEs) are yet to reap these benefits as obtaining such opportunities rests on their ability to integrate ICT into business tasks and processes. They need to transform their companies by implementing appropriate ICT in order to compete within a busy digital economy. The aim of this study was to see how SMEs from developed and developing countries follow their eTransformation journey and also to identify the differences and similarities businesses encounter in that process. Our research confirms that the eTransformation takes place as an incremental process for both groups and that the eTransformation pattern is similar in both worlds. However, we notice that in Australia, a developed country, SMEs are more aware of eTransformation and are demanded by customers and digital economy to change, where as in Sri Lanka eTransformation is still in its early stages of development.

1   INTRODUCTION

Over the last few decades the world economy has undergone numerous changes and has been impacted by rapid advances in technology. The knowledge demand on how to access markets and business partners or suppliers worldwide with the aim to remain competitive strongly depends on the company's ability to eTransform, the process through which businessess identify, select and implement required technology (Hol & Ginige, 2009).

According to Ginige and colleagues (2001) the process of eTransformation is seen as a staged and incremental. There are a number of models of eTransformation, however there are two that explain the process in detail. One of the early models is the eTransformation Road Map model (Ginige et al, 2001). This model has been widely used to explain eTransformation process within developed countries (Lawson et al 2007; Hol and Ginige, 2008, 2009). The eTransformation Road Map has been extended further into eSME model by the researchers within the developing countries (Kapurubandara, 2009b). Both models stress that in order to eTransform successfully SMEs are required to complete eTransformation steps incrementally.

SMEs play a major role in many economies, globally, irrespective of the fact whether they are in developed or developing countries. They contribute significantly to the economies of many countries of the developed world (Rao, 2003). In Australia, the SMEs are not only considered major employers (70% of Australian employment), but they are also viewed as being major drivers of innovation, and as such, key players in making an economy competitive in the global marketplace (Temperley, 2004).

In Sri Lanka 75% of the labour force is employed in the SME sector (including the agricultural sector) (ILO, 2006). These numbers suggest that SMEs contribute substantially to employment and income generation.

SMEs in developing countries face an abundance of external constraints due to the differences in infrastructure, social, cultural and regulatory factors and internal constraints inherent to SMEs by nature (Kapurubandara et.al, 2007).

This effect is also visible within developed SMEs in Australia. In developed countries, SMEs seem to be aware of the technology but still hesitant, lacking time and resources to fully embrace it (Lawson et. al. 2007).

This study aims to identify similarities and differences of SMEs eTransformation within developed and developing worlds. The comparative investigation was carried using two selected groups of 3 SMEs each, one from developing and the other from the developed world.

2   METHODOLOGY

To identify how SMEs within developing and developed countries, in this case Sri Lanka's Colombo District and Australia's Sydney metropolitan region are selecting, implementing and using technology to speed up and complete their business tasks and processes quicker and swifter we chose an exploratory study.

The subjects of our study were 3 SMEs from a developing country – Sri Lanka's – Colombo region and another 3 from a developed country – Australia's Sydney region. The sample for Sri Lanka was chosen from a list maintained by the Export Development Board of Sri Lanka a representative of SMEs from various industry sectors, within Colombo. The sample for Australia was chosen from a list of previously studied sample of 30 companies belonging to a range of industry sectors selected from Corporate Development Office.

To conduct the study CEOs or Directing Managers of SMEs were interviewed to get first hand information about the business tasks and processes, their use of ICT, and their perceptions for adopting it.

Each interview lasted nearly one hour. Interviews were recorded, subsequently transcribed and subjected to qualitative content analysis. This involved categorisation of responses into certain themes, some of which were inherent in the interview questions, and some of which emerged through the interviewees' responses to questions posed.

3   ANALYSIS AND DISCUSSION

Data collected during the interviews was analysed and categorised. Sri Lankan based companies ranged in size from 6 to 12 employees. SMEs used fax, phone and generally performed record keeping manually. Two out of three companies had stand alone computers while one had a network and ADSL.

In general, computers are not fully utilised in a cost beneficial manner, but more as a storage medium. SMEs are lacking in resources, know how, awareness of the optimum usage of ICT. They face more burning problems and are concerned with day to day survival; hence ICT implementataion has taken less priority.

Selected Australian based SMEs ranged in size from 3 to 54 employees. They all had networks, were continuously contacted to the Internet and all used productivity tools. All records were kept in the electronic form however they lacked searching, managing and mass customisation tools. SMEs in Australia all use productivity tools and specialized industry software however are lacking systems such as CRM (Customer Relationship Management), tracking systems, document management systems and interactive online systems.

In addition, to better understand process of eTransformation we mapped all six SMEs on then eSME Roadmap (Kapurubandara, 2009b). Data is presented in Table 1.

Table 1: Internal and External Processes on eSME.

External Processes	SME (SL)	SME (AU)	Internal Processes	SME (SL)	SME (AU)
Stage 1 No Website	S1, S2, S3		Manual	S1	
Stage 2 Basic Website			Effective Individual	S2, S3	
Stage 3 Interactive Website		A1, A2, A3	Effective Team		A1, A2, A3
Stage 4 E-Commerce Site			Effective Enterprise		

Table 2: Similarities and differences between the two groups of SMEs.

SMEs (SL)	SMEs (AU)
Web Not an essential tool for marketing. Word of mouth is used for marketing	Websites are present, however main jobs come via word of mouth.
email is used occasionally but not the preferred medium of communication	e-mails aare used regularly. Companies are trying to establish Docoment Management systems for easy file storing and accessing.
Lack skills, resource, knowledge of effective use of ICT.	Lack of skills, knowledg and ICT resources is still present, however most SMEs have realised this and are trying to make changes. They are aware of eTransformation.
eTransformation is given less priority.	eTransformation is seen as important, but essential only when companies realise that old methods do not work any more.
Manual processes are preferred. However computers are used as storage medium	Manual procesess are still popular, and filing is done in both paper copies and electroniclly.
Customers have not adopted ICT. Lack of pressure from the customers.	Customers are demanding change as they search for cheaper and better products.
Not much attention is given for quality. SMEs need to be educated	Quality of the product as well as timely delivery are becoming a priority.
SMEs are aware of the importance of customers. Lack funds for implemantaation of CRM. Awareness need to be created in systems such as CRM.	Companies are identifying that the customers are their crucial asset. CRMs are becoming popular.

Furthermore, it also became evident that SMEs can be in different levels of maturity with respect to their internal and external processes which is in common with the other stages of growth models (McKay, Prananto, & Marshall, 2000).

Our investigation indicates that Sri Lankan SMEs broadly fit into stage 1 and stage 2 on the internal process arm of the eSME roadmap. SME S2 and S3 fall into two different sub stages with in the stage "Effective Individual". In Sri Lanka, e-mail and Internet are considered costly affairs (SLBDC, 2002). Where the SMEs are saddled with many other pressing concerns, adoption of eTransformation given a lower priority which is not surprising given the expenditure involved (Lawson, Alcock, Cooper, & Burgess, 2003; Lawson et. al. 2005).

On the other hand, Australian SMEs have reached Interactive website (Stage 3 on the external processes) and effective team (they are between stage 2 and 3 of the Internal Processes) which indicates that they have advanced further on the eSME Roadmap.

It can be concluded that SMEs lack awareness of eTransformation and are struggling with costs associated with hardware/software purchase which led to only partial adoption. Previous studies confirm that lack of awareness can significantly impede technology adoption (Bode, & Burn, 2001).

Table 2 surmises similarities and differences between SMEs from developed and developing countries.

This exploratory study has shown that while the two groups of SMEs agree that adoption is essential in today's business they are hindered from adoption. This is predominant with the Sri Lankan SMEs, but still present within Australian SMEs too. When compared to the Australian SMEs, Sri Lankan SMEs lack in awareness, resources, knowledge and skills required for a successful eTransformation while Australian SMEs are aware of eTransformation but still do not have required skills to fully implement it. This could be attributed to the fact that eTarnsformation was introduced in Australia much earlier than Sri Lanka and that businesses within Australia have a slight advantage of early adopters. Furthermore, it appears that some of the companies have been forced to implement ICT as they have realised that their existing processes were not functioning any longer, which may happen within Sri Lanka as well in the years to come (Lawson et. al. 2005).

When analysing the process of eTransformation however, it seems that both groups follow a similar pattern in their eTransformation journey irrespective of the fact whether they are from a developing or a developed country. It looks that their journeys have started at different times and also that their rate of progression is different. Australian companies are ahead, however they have been faced in the past with very similar challenges with which Sri Lankan companies are faced with today. We can also see that eTransformation journey is staged and incremental. In addition, a more detailed eSME

Roadmap depicts this journey in minute steps and identifies stages SMEs are required to undertake for both developed and developing countries.

3 CONCLUSIONS

In summary it can be seen that within both developing – Sri Lanka and developed – Australia countries SMEs today are required to undertake the process of eTransformation. By analysing eTransforming SMEs it can be concluded that SMEs are required to undertake an incremental, staged journey in order to eTransform. However due to many inherent constraints, the SMEs in developed countries are forced to take minute steps in discovering eTransformation and basic ICT tools. On the other hand, the SMEs in developed countries now have an understanding of what eTransformation is and are on the road to discovery of what eTarnsformation and ICT system integration can offer to them. Due to the governmental regulations and country's development and access to the technology Australian SMEs have undertaken the eTransformation journey before Sri Lankan SMEs; however the journey Sri Lankan SMEs are undertaking today is very similar to the journey Australian SMEs have been undertaking earlier. Furthermore, pressures of the digital worlds, government and customer demands have also contributed to swifter eTransformation within developed countries. It is expected that in the years to come Sri Lankan SMEs will be undertaking a similar journey Australian SMEs are undertaking today.

REFERENCES

Bode, S., & Burn, J. M. (2001). Website design consultants, Australian SMEs and electronic commerce success factors. *International Journal of Business Studies*, 9(1), 73-85

Ginige, A., S. Murugesan, (2001). "A road map for successfully transforming SMEs into e-businesses." *Cutter IT Journal* 14(5): 43.

Hol, A. and Ginige, A. (2009) eTransformation Journey for SMEs, *ICE-B International Conference on eBusiness*

ILO. (2006). Improving the local business environment through dialogue; a story from Sri Lanka.

Kapurubandara, M., (2009a). A Framework to e-Transform SMEs in Developing countries, *The Electronic Journal of Information Systems in Developing Countries* (EJISDC), Vol 39, ISSN: 1681-4835, www.ejisdc.org

Kapurubandara, M. (2009b). eTransformation of SMEs. A Framework for Developing Countries. Germany, Vdm Group Inc ISBN: 978-3-639-21971-5;

Kapurubandara, M., & Lawson, R. (2007). SMEs in Developing Countries face challenges in adopting e-commerce technologies. Paper presented at the *IEEE DEST Conference Series* 2007, Cairns, Australia, February 2007, ISBN: 1-4244-0470-3, pp 141-146 46, Digital Object Identifier: 10.1109/DEST.2007.

Lawson, R., C. Alcock, et al. (2003). "Factors affecting adoption of electronic commerce technologies by SMEs." *Journal of Small Business and Enterprise Development* 10(2): 265-276.

Lawson, R., Arunatileka, S., Ginige, A. and Hol, A (2005), "A Pilot Project on eCollaboration in the Australian Toolmaking Industry", in *Proceedings of the 18th Bled eConference eIntegration in Action International Conference, Bled, Slovenia*, June 6 - 8, 2005.

Lawson, R., Hol, A. and Hall, T. (2007) Challenges of eCollaboration among SMEs, *Proceedings of the 20th Bled eConference, eMergence: Merging and Emerging Technologies, Process and Institutions*, 4-6 June, 2007 Bled, Slovenia

McKay, J., Prananto, A., & Marshall, P. (2000). E-business maturity: the SOG-e model. Paper presented at the *Australian Conference on Information System (ACIS)*, Brisbane, Australia.

Rao, S. S., Metts, G., & Monge, C. A. M. (2003). Electronic commerce development in small and medium sized enterprises: A stage model and its implications. *Business Process Management Journal*, 9(1), 11-32.

Temperley, N. C., Galloway, J., & Liston, J. (2004). SMEs in Australia's High - Technology Sector: Challenges and Opportunities CSIRO/AEEMA. Retrieved 22 July 2006 from http://www.aeema.asn.au

FUZZY DECISION MAKING OF IT GOVERNANCE

Sung Eui Cho, Sang Hyun Lee

Department of Computer Engineering, Mokpo National University, Jeonnam, Korea
leesang64@gmail.com, chosug@mokpo.ac.kr

Kyung-Il Moon

Department of Computer Engineering, Honam University, Gwangju, Korea
kimoon@honam.ac.kr

Abstract: IT governance implies a system in which all stakeholders, including the board, internal customers and related areas such as finance, have the necessary input into the decision making process. IT governance is the preparation for, making of and implementation of IT-related decisions regarding goals, processes, people and technology on a tactical or strategic level. But, the concepts of IT governance are broad and ambiguous which in turn implicate difficult and inaccurate assessments. In particular, the traditional handling of IT management by board-level executives is that due to limited technical experience and IT complexity, key decisions are deferred to IT professionals. This paper presents a fuzzy reasoning model for assessing IT governance complexity based on an extensive literature study. This model can be used for a good understanding how the concerns of IT governance behave, how they interact and form the behaviour of the whole system. The model for assessing IT governance is employed to compare how IT governance is defined in practitioners and Cobit.

1 INTRODUCTION

This paper suggests a collective behaviour model based on fuzzy reasoning with respect to IT governance concerns considered important in literature, and to represent how the concerns should be really addressed by practitioners and Cobit. Understanding how the concerns of IT governance behave, how they interact and form the behaviour of the whole system can certainly be interesting through this model. Factors such as concerns (and the number of them), interaction between concerns, environment, and IT governance activities can be equally interesting when studying "self-organized" IT governance systems, if the aim is understanding. But when we go about designing, a control system, we will be guiding its organization and we need to understand the complexity of the concerns and the emerging whole. Returning to complex interaction, we feel a need to attempt relating the system.

2 IT GOVERNANCE COMPLEXITY PROFILES

Complex system typically has some characteristic properties, but the extent to which a particular system exhibits any given property can vary. In this respect, IT governance complexity is a nonlinear mapping concept.

2.1 Domain Complexity

The domain complexity denotes a nonlinear function of what the decisions should consider. It comprises four complexity variables: goal, processes, people and technology. Goals include strategy-related decisions, development and refinement of IT policies and guidelines, and control objectives used for performance assessments. Processes include the implementation and management of IT processes, e.g. acquisition, service level management, and incident management. People include the relational architecture within the organization, and the roles and responsibilities of different stakeholders. Finally, IT governance is of course about managing the

technology itself. The complexity variable technology represents the physical assets that the decisions consider, such as the actual hardware, software and facilities.

2.2 Scope Complexity

The scope complexity denotes a nonlinear function of different impacts implied by each decision. There is a long term aspect and a short time aspect of every decision that is made. The scope dimension is used to differentiate between different levels of decision-making. Firstly, there are detailed, rapidly carried out, IT-focused tactic decisions. Examples of tactic decisions include whether to upgrade a certain workstation today or tomorrow, how to configure a user interface that is only used internally, or the manning of a single IT project.

There also exists top management, low detailed, business oriented strategic decisions with long timeline. A strategic decision might consider whether it is most appropriate to develop an application in-house or to purchase it off the shelf, or how the performance of IT processes should be reported to top management.

2.3 Decision Making Complexity

The decision complexity denotes a nonlinear function of different steps required to make decisions within the different domains. This complexity deals with the relation between IT, and the models of the reality used for decision making. Before making any decision regarding e.g. the outsourcing of a helpdesk function, the organization must be clearly understood. Facts have to be thought over and investigated, and transformed into a model. The model might be a simple cognitive map, present nowhere else but in the head of the decision-maker, or a more formalized, abstract model put on print.

This process of analysis and understanding is denoted the understanding phase. Once the model is created, the actual decision can be made according to corporate IT principles, in a timely manner, by the right individuals, etc. In the IT governance definition, this is represented by the decision phase, which also includes planning of how to make the decision.

3 IT GOVERNANCE COMPLEXITY MODEL

The objective of this section is to understand the relationship between the complexity profiles and to construct a fuzzy reasoning model including the complexity of collective behaviour with respect to IT governance.

3.1 Relationship between Complexity Profiles

IT governance is not strict hierarchy. It contains lateral interactions that enable control to bypass the hierarchy. However, by focusing on an idealized control hierarchy it is possible to understand the nature of this structure. Such a focus will help in understanding the relationship between this structure and complex collective behavior. In an idealized hierarchy all communication, and thus coordination of activities, is performed through the hierarchy. Figure 1 denotes a hierarchical network structure. It describes the content of different statements identified in literature with respect to IT governance concerns.
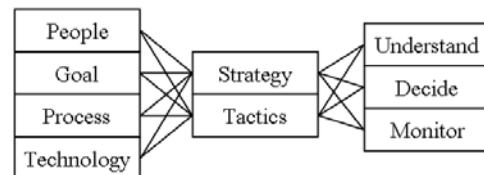
Figure 1: Hierarchical Structure of IT governance.

This structure imposes a limitation (say, network weight) on the degree of collective behaviours of IT governance. This can be understood by considering more carefully the processes of coordination. The hierarchy is responsible for ensuring coordination of various concerns of IT governance. Lower levels of the hierarchy are responsible for locally coordinating smaller parts of IT governance and higher levels of the hierarchy are responsible for coordinating the larger parts of IT governance.

3.2 Fuzzy Reasoning Approach to Complexity Profiles

Fuzzy rules are usually formulated as IF-THEN statements, with one or more antecedents connected to a consequent via operators like AND, OR, etc.

IF ($Antecedent_1$) OP ($Antecedent_2$) … OP ($Antecedent_n$) THEN (Consequent) (w)

Here n is an integer, OP is standing for operators like AND, OR, etc., and w represents a weight value indicating the importance of a rule. In this study, our

fuzzy reasoning method is based on two assumptions as the following.

■ Every activation of an input fuzzy set is regarded to be a piece of (fuzzy) concerns supporting the domain knowledge an expert formulated via rules and fuzzy sets.

■ Each piece of concerns should be incorporated more actively in the decision-making process.

Table 1: Example of 3 steps.

	DM		
	low	*normal*	*High*
Domain	0.86	0.37	
Scope		0.77	
Accumulation	0.86	1.14	
Normalization	0.75	1.00	
Decision making	DM=DM$_{normal}$/1.00		

These assumptions can be implemented in 3 steps, concerns accumulation, normalization, decision-making. For example, Table 1 illustrates an application of 3 steps. The accumulation of the pieces of concerns produces: DM$_{low}$=0.86, and DM$_{normal}$=0.37+0.77= 1.14. Normalization of these values generates: DM$_{low}$=0.75, and DM$_{normal}$=1.00. The method therefore produces the outcome: DM=*normal*. This approach can be also applied to aggregation of the consequents across the rules, as there are many different weights indicating the importance of the rule.

When available computational capabilities are restricted by equipment size or cost, special attention should be given to defuzzification process. In these cases, the computational time must be reduced in order to improve the system performance. Hence, it is important to use fast defuzzification methods. As an alternative, faster and simple methods can be used such as finding the mean of maxima or by finding the half-area point. we presents a simple fast method for computing a centroid approximation by fitting the fuzzy output area into a triangular shape, see figure 2.

This approach consists in adapting any output shape into one single triangle. The computational time required by this algorithm is reduced with respect to that of the bisector method. This approximation gives the exact centroid position for any cluster shapes having a base length and areas ratio of 1 to 3. For fuzzy outputs not located at the origin, the triangular shape maximum position is located at the maximum output shape position. When the fuzzy output presents more than one

maximum, the location of the triangle maximum is computed as the average of maxima.
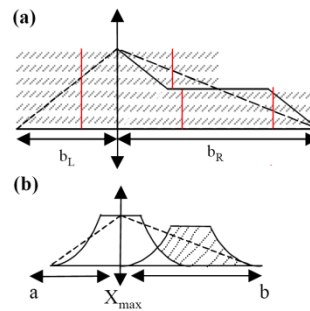
Figure 2: Output fitting to triangular function.

4 APPLICATION

Table 2 shows the results for these theoretical complexity variables, i.e. literature's concerns of IT governance. The total score for each dimension is 100%. Also, it includes the normalization within each dimension complexity, explained in previous section. Related to Domain complexity, twelve rules are defined in the rule base. We used the normalized rule weights for fuzzy pieces of IT governance concerns where many rules apply to the same conclusion, and used the simple fast defuzzification method in the previous section.

Table 2: IT Governance concerns according to literature.

Dim.	Complexity Variables	Literature Concerns	Normal
Domain	People	0.37	1.00
	Goal	0.26	0.70
	Process	0.20	0.54
	Technology	0.17	0.46
Scope	Strategy	0.70	1.00
	Tactics	0.30	0.43
DM phase	Monitor	0.42	1.00
	Decide	0.33	0.79
	Understanding	0.25	0.60

The theoretical IT governance concerns show that the dimensional variables "People", "Strategic", and "Monitor" were most frequently used within the 50 articles and within their dimensions respectively. IT governance mainly comprises strategic concerns according to literature. The daily use of IT, all the operational concerns for bread-and-butter IT are surely important, but they are not in the scope of IT governance. Regarding the decision-making phases, monitoring of IT-related decisions is emphasized. Technology issues are not the mayor concerns to

decide upon, and literature rather stresses the importance of establishing roles and responsibilities, and an accountability framework that supports the business goals.

In the Fuzzy DM of IT governance concerns, there are five parts of the fuzzy DM process: fuzzification of the input variables, treatment of the fuzzy pieces evidence, implication from the antecedent to the consequent, aggregation of the consequents across the rules, and defuzzification. Figure 3 only illustrates fuzzy sets for Domain complexity. The two variables have each been divided into 3 overlapping sets labeled Low, Normal, and High. The first vertical line represents a measurement of Process, which has a membership level of 0.2 in all the Low, Normal, and High sets, c.f. (Table 2). The second represents a measurement of Goal, which has a membership level of 0.26 in all sets. We can construct fuzzy sets of Scope and DM complexity in a similar manner. Related to Domain complexity, 12 rules are defined in the rule base.
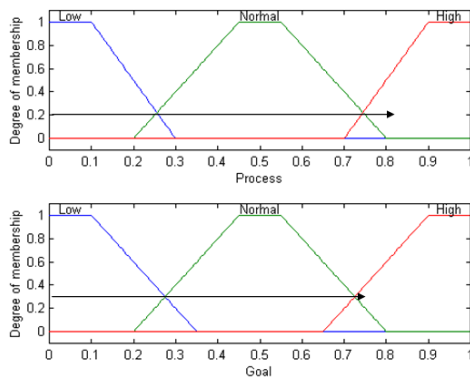
Figure 3: Fuzzy sets for Process and Goal in Domain complexity.

Figure 4 shows the surface plot between input variables of Domain complexity. Clearly it is evident from the plot that "People" is more significant than other input variables. IT governance concerns in Literature denotes that "Technology" is less significant than other ones. But, considered as a whole, "Process" is less significant than other ones, c.f. (Table 3). In particular, in proportion as "Goal" rises "Technology" concerns increase. Table 3 illustrates the comparison of values estimated by using four input variables.

According to the survey with practitioners, practitioner's concerns were mainly about IT goal setting, while IT processes and technology issues were less stressed. Table 4 illustrates the comparison of values estimated by our fuzzy model. Here, "Goal" is more stressed. Table 5 illustrates the
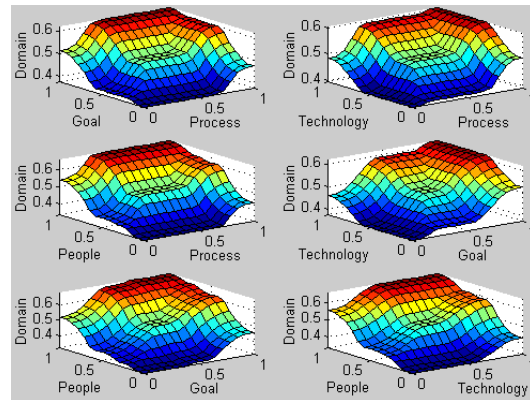
Figure 4: Mapping surface of Domain complexity.

Table 3: Comparison of values by fuzzy model.

Process	Goal	Technology	People	Domain
0.2	0.4	0.6	0.8	0.571
0.2	0.6	0.4	0.8	0.571
0.4	0.6	0.2	0.8	0.564
0.4	0.2	0.6	0.8	0.527
0.6	0.4	0.2	0.8	0.564
0.6	0.2	0.4	0.8	0.527

Table 4: Comparison of values by practitioners' concerns.

Process	Goal	Technology	People	Domain
0.2	0.8	0.4	0.6	0.536
0.2	0.8	0.6	0.4	0.536
0.4	0.8	0.2	0.6	0.527
0.6	0.8	0.2	0.4	0.527

comparison of values estimated by our fuzzy model. The result denotes that there is discrepancy in the range of the concerns identified in literature. Figure 5 shows the surface plots between input variables of DM and scope complexity, respectively. For DM complexity, the nine rules and normalized weights are included in the fuzzy rule system.

The theoretical concerns showed that the dimensional variable "Monitor" was more frequently used within the DM complexity. But, monitoring the implementation of decisions already made receives somewhat less attention from the practitioners, according to the survey. Also, comparing Cobit's concerns of IT governance to literature, it showed that Cobit does support most needs, but lacks in providing information on how decision-making structures should be implemented.

Table 5: Comparison of values by Cobit.

Process	Goal	Technology	People	Domain
0.8	0.6	0.2	0.4	0.492
0.8	0.4	0.2	0.6	0.492
0.8	0.2	0.4	0.6	0.460
0.8	0.2	0.6	0.4	0.460

Applied to our fuzzy model, the dimension variables of DM complexity are almost uniformly stressed. The relative concerns for the DM complexity remain a bit more uncertain. The difference seems to lie in their interconnection weights (and interactions) between the concerns of IT governance. For scope complexity, strategic concerns are most often dealt with, while tactical concerns are only briefly discussed. The six rules and normalized weights are included in the fuzzy rule system.

IT governance mainly comprises strategic concerns according to literature. According to the practitioners responding the survey, IT governance decision making is mainly a strategy issue while tactical decisions are less important. Similarly, Cobit spends more effort in discussing strategic concerns and less on tactical concerns. But, according to the mapping surface of Figure 5, strategic and tactical concerns that make up a large collective behaviour must be correlated and not independent.
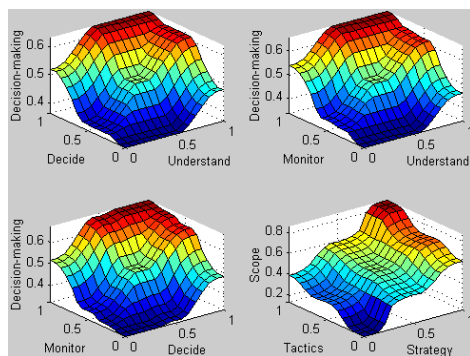
Figure 5: Mapping surface of DM and Scope complexity.

5 CONCLUSIONS

This paper presented a framework to understand the relationship between the complexity profiles in view of complexity science, and then developed a fuzzy reasoning model including the complexity of collective behaviour with respect to IT governance. It is necessary to understand the exact nature of the interconnections and how their weights give some effects on the behaviour of the whole IT governance. When there are such interconnections and they are not simple, a complex system can be used. In particular, IT governance complexity is a fuzzy concept. Thus, we suggested a fuzzy model for analyzing IT governance complexity based on an extensive literature study. IT governance concerns in literature were mapped onto the framework for this

model, and a comparison study was carried out. Results showed that the major differences exist within the concerns of the domain complexity in the case of Cobit.

REFERENCES

Dahlberg, T., and Kivijärvi, H., 2006. An Integrated Framework for IT Governance and the Development and Validation of an Assessment Instrument. *Proceedings of the 39th Hawaii International Conference on System Sciences.*

De Haes, S., and Van Grembergen, W., 2005. IT Governance Structures, Processes and Relational Mechanisms – achieving IT/Business alignment in a major Belgian financial group. *Proceedings of the 38th Hawaii International Conference on system Sciences.*

Holm Larsen, M., Kühn Pedersen, M., and Viborg Andersen, K., 2006. IT Governance – Reviewing 17 IT Governance Tools and Analysing the Case of Novozymes A/S. *Proceedings of the 39th Hawaii International Conference on System Sciences.*

Johansson, E., 2005. Assessment of Enterprise Information Security – How to make it Credible and Efficient. *Ph.D. Thesis at the Department of Industrial Information and Control Systems, Royal Institute of Technology*, Stockholm, Sweden.

Ridley, G., et al., 2004. COBIT and its utilization: A framework from the literature. *Proceedings of the 37th Hawaii International Conference on System Sciences.*

Wang, H. F., and Huang, Z. H., 2002. Top-down fuzzy decision making with partial preference information. *Fuzzy Optimization and Decision Making 1*, 161-176.

Warland, C., and Ridley, G., 2005. Awareness of IT control frameworks in an Australian state government: A qualitative case study.

Webb, P., Pollard, C., and Ridley, G., 2006. Attempting to define IT Governance: Wisdom or Folly. *Proceedings of the 39th Hawaii International Conference on system Sciences.*

Yager, R. R., 2002. On the evaluation of uncertain courses of action. *Fuzzy Optimization and Decision Making 1, 13-41.*

ENTERPRISE ENGINEERING

FULL PAPER

A FRAMEWORK TO DERIVE HOLISTIC BUSINESS TRANSFORMATION PROCESSES

Bhuvan Unhelkar and Athula Ginige

AeIMS Research Group, University of Western Sydney, Sydney, Australia
Bhuvan.unhelkar@gmail.com, a.ginige@uws.edu.au

Abstract: This paper describes an approach to deriving a holistic business transformation process and its application in practice. When a business is transforming there needs to be changes to many aspects of its Enterprise Architecture (EA). The research has ascertained that often organisations are not aware of all the aspects that need to be changed to successfully implement a business transformation. This paper presents a holistic approach to transformation that identifies the current EA of the organisation and models the one that should be after transformation to identify all aspects that need to under go transformation. It then uses a meta-model of a process to derive the required processes to achieve a holistic transformation of the businesses undertaking electronic transformations. This approach evolved in the process of assisting business transformation is the Small and Medium Enterprise (SME) sector within the Western Sydney region. This is demonstrated through a real-life example. This work has been carried out under the auspices of the AeIMS Research Group at the University of Western Sydney.

1 INTRODUCTION

Rapid advances in Information and Communications Technologies (ICT) imply a change in the way businesses are organized and operated. The concept of agility in business, which is a result of ICT innovations and advances, leads to collaborative and competitive global transactions(Unhelkar et al., 2009). The need to understand, model, review and implement efficient yet dynamic business processes could not be overemphasized. However, there also appears to be a dearth of appropriate holistic transformation processes that can be used for modeling and transforming business processes.

This paper aims to fill that lacuna by deriving the process of business transformation. Furthermore, this paper also outlines the application of this approach to real life business transformation projects. The factors influencing business transformations and the risks and advantages associated with them have been reported by Arunatileka S. and Ginige (Arunatileka and Ginige, 2003b). These factors have also been identified and extended by Arunatileka , D., et.al(Arunatileka et al., 2008) and applied to mobile businesses. Electronic

transformations of businesses aim to capitalize on the connectivity accorded by the ubiquitous Internet. These transformations result in numerous business benefits such as enhanced customer experience and improved internal business efficiency(Deitel et al., 2001) (listed later in this paper in detail). However, business transformations themselves need a well thought out process that helps identify the business goals, the current structure of the business and steps in undertaking the necessary change to become an electronic-global business. Enterprise Architecture (EA) frameworks (such as Zachman, TOGAF) (The_Open_Group, 2010, Zachman and Holcman, 2010) provide excellent backdrop for the study of enterprises. We extend and modify these frameworks in understanding the business as it stands. However, our core motivation of this work is the way in which these frameworks can be used to model the future expected state of the business and how to reach there. Consider, for example, the 6 x 6 model of a business as presented in the Zachman framework. We abstract the building blocks of a business from this EA framework and summarize it into technology, networks, data, functions/processes and people. Each of these elements can be made up of many sub-elements depending on the type and

size of business and its transformation goals. We argue that modeling these elements along the lines we have discussed later in this article, is a crucial step in undertaking transformation.

The rest of this paper is organized as follows: project background and literature review, dimensions of a transformation framework, derivation of a business transformation process, discussion on the approach to validation of the transformation framework based on application, conclusions and future directions.

2 LITERATURE REVIEW & PROJECT BACKGROUND

Precursor to this work is the earlier reported work on business transformations that was undertaken by Ginige et al (Ginige, 2006, Ginige, 2008, Hol and Ginige, 2008) and Ghanbary and Unhelkar (Ghanbary and Unhelkar, 2007). These earlier reports underscored the fact that formal planning and execution was vital in undertaking successful business transformations. The advances of semantic web and service orientation have further propelled the rate at which the business environment is changing as we move away from information silos to collaborative processes. The rapid impact of IT demands derivation of a formal and systematic transformation process. Use of such process will lend itself to configuration depending on the type, size and goals of the organization. Numerous factors influence these transformations – including the economic, technical, process and social dimensions(Hol and Ginige, 2008, Unhelkar, 2010).
Figure 1 (A) explains why businesses undertake formal transformations. The gap between where the organization is and its external environment presents a continuous challenge to the business. The external environment acts on the organization which, in turn, takes its time and responds with changes to its processes, technologies and people. The time taken between the initial action of the environment and the response of the organization keeps it "out of sync" with the external environment. Businesses want to transform themselves in order to become agile – so that they are able to respond effectively to changing external circumstances and stay in harmony with it. Figure 1 (B) summarizes what the organization needs to do in order to transform: have a business transformation process framework, derive and configure a specific transformation process from it, execute the process by enacting it and finally maintain the transformed operational processes on

an ongoing basis. Such ongoing maintenance is what constitutes an agile organization that remains in harmony with the external environment.

We are understanding agility of business as a time measure between two significant changes in the environment and time it takes for an organisation to respond to that change. A highly agile organization will be able to quickly respond to a change in the environment; however, as the rate of environmental change increases the corresponding agility of the organization can become low. At some point the need for an organization to keep up its responses to changing business environment becomes a continuous process. This is what a truly agile organization is – and it provides another reason for formal transformation processes.

Following are the specific aims of an electronically transformed agile business:

(a) Extend its reach globally and provide wide coverage for the organization's products and services with the use of communications technologies and web services,

(b) Change the business processes of the organization to quickly and effectively responding to changing needs of the customer in a location and time independent manner,

(c) Enhance customer experience through personalized services and additional services,

(d) Reduce internal costs and improve the quality and efficiency of its procurement, inventories and development activities ,

(e) Improve corporate accountability and regulatory compliance through timely, accurate and detailed reporting on business performance and

(f) Manage environmental and socio-cultural responsibilities with efficient business processes.

These advantages of business transformation and globalization have also been studied by Lan and Unhelkar(Lan and Unhelkar, 2005) under Global Enterprise Transitions, and separately outlined as seven principles of electronic transformation (e-Transformation) by Arunatileka et al (Arunatileka and Ginige, 2003b, Arunatileka and Ginige, 2003a). While earlier attempts to undertake business transformations were radical (such as reported by Hammer and Champy (Hammer and Champy, 2001) in re-engineering the corporation and also electronic transformations roadmaps discussed by Kalakota and Robinson (Kalakota and Robinson, 2001), our experiences – especially in the context of SMEs – were that such radical transformations did not succeed and that their stated business benefits did not accrue.
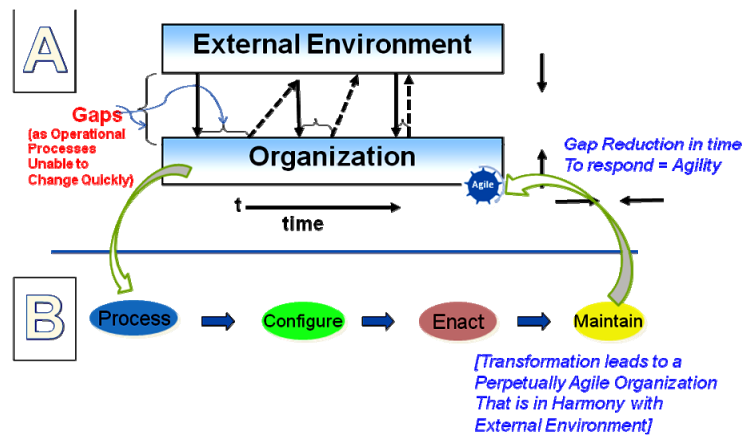
Figure 1: Goals of Business Transformations is to produce an Agile Organization that has minimal time gap between action from the external environment and response of the organization.

Our experiences and further investigations suggest that transformation of business requires careful handling of the operational processes of the organization. This focus on operational business processes in transformations is further ratified by the work of Arunatileka D.(Arunathilaka, 2006) who investigated an in-depth approach to transformation of mobile business processes. In a recent global survey of practitioners, Unhelkar (Unhelkar, 2009) also discovered that business perceive "business processes" as a major area of value as well as risk. Hence a transformation process needs to pay particular attention to the operational business processes.

Based on our literature review, understanding and experiences with SMEs, we find that the success of business transformation depends on the following factors:

- The purpose of transformation and business goals need to be clearly spelled out. These goals, as mentioned earlier, can range from cost reduction to enhancing customer experience. More importantly, these goals can change depending on the type, size and strategic direction of the organization.
- Understanding of a well formulated business transformation process. Deriving a transformation process formally from an existing framework is a vital part of successful transformation. Not having such a process is a risk to the business wanting to change. The value of a transformation process is its ability to configure itself.
- Enterprise Architecture frameworks help in modeling the existing and targeted organization. Such EA frameworks provide, in measurable

terms, what gets transformed and what value it provides to the business. They also provide the basis for deliverables used in mapping the existing and future organization.

- Modeling of the operational processes of the organization is critical in undertaking successful transformation. As mentioned earlier, operational business processes provide the single most critical aspect of a business and they need to be modeled, transformed and validated for the success of the new business.
- IT systems and applications need to be studied and their migration needs to be planned out in a phased manner. The organization continues to function while it undergoes transformation. Therefore there is a need to support the organization while it operates and, at the same time, transform its applications.
- Content Management systems – especially dealing with rapidly changing multimedia contents – need to be modeled and a road map for their transformation outlined. Business transformation will require sourcing of new contents, correlation of existing contents with each other, and cohesive display of location-specific contents to users.
- Communications and networks also undergo change – especially for large and global organizations. Transformation of communications and networks is a specialist technical dimension of the transformation and needs to be factored in the overall transformation.
- Quality assurance of the new enterprise including verification and validation of the operational processes that have been transformed.

3 BUSINESS TRANSFORMATION FRAMEWORK

Figure 2 presents a generic meta-model of a process. This meta-model can be used to derive a specific process of transformation. Usually, this derivation is dictated by the goals of transformation and the subsequent deliverables. For example, a small transport company would undertake transformation of its operational processes only, whereas a large auto manufacturer would undertake transformation of all dimensions of its business. There is further need to segregate processes that relate directly to the operation of the business and the process that handles the business transformation. Unhelkar(Unhelkar, 2003) has argued for three core elements of a process that describes what gets transformed, who does it and how it is done – the deliverables, roles and activities respectively. Against the backdrop of these arguments, we propose core elements of the meta-model for the business transformation process (Figure 2). Successful electronic transformation needs to consider the activities and tasks that need to be undertaken for transformation, the people who will undertake the transformation and the deliverables that will be produced at the end of the exercise.

These core elements of a business transformation meta-model can be further described as follows:

A suite of deliverables; these are the templates that help describe the existing enterprise and also the transformed enterprise. The deliverables based on the templates reduce the risks in undertaking transformation and improve its quality. For example, mapping the organization as containing data, functions and people requires each of these elements to be described. Templates providing that description are a part of the transformation meta-model and can be used for this purpose. Our experience suggests that a practical transformation process is iteratively derived based on the required deliverables and business goals. For example, in case of Small and Medium Enterprises (SMEs), not all types of deliverables– such as user training or package implementation - are required. In case of large and global organizations, on the other hand, all elements of deliverables are required including their lifecycles and iterative updates.

A suite of activities and tasks; these are the step-by-step guidance provided by the process for transformation. For example, identification of the goals of transformation, documenting the operational functions/processes and applying verification and validation techniques to the deliverables are all listed as formal activities within the transformation process.

Roles that undertake transformation and those that get transformed; the people involved in transforming the business include the decision makers, business analysts, testers and regulators. The roles that undergo transformation include management, external parties and customers. A transformation meta-model contains all possible roles from which some are instantiated depending on the actual transformation taking place.

4 DERIVING A BUSINESS TRANSFORMATION PROCESS

There are many models to represent different aspects of an organisation. 7S is one such model. 7e in eTransformation model was derived based on 7S model (Arunatileka and Ginige, 2003b) Thus, eventually, as explained earlier in Figure 1 (B), through a transformation process we aim to create an agile organization that is "in sync" with its external environment. The model to understand an organization selected by us is the Zachman model. The overall process for deriving a business transformation is as follows:

- Note the type and size of the organization – small, medium and large organizations would transform differently
- Identify the status of the current organization using a known model, such as Zachman model
- Identify the goals of transformation of the organization in short and long term
- Identify and document strategies to respond to the changes in the business environment.
- Identify the affected processes and approach required to change these business processes.
- Use Zachman model to identify what should be in rows and columns to implement these new processes.
- Identify all cells in Zachman where there is a difference between existing elements and desired ones
- Use Process meta-model presented earlier to derive a Transformation processes for each pair of related cells that are different.
- Based on the difference formulate the deliverables for each aspect that needs to change.
- Based on the deliverables derived each of the components of Transformation process.
- Work out the detail activities required.

144
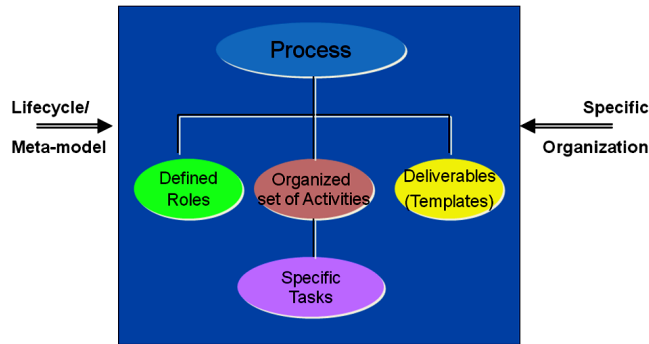
Figure 2: Core Elements of a Process Meta-model that can be used to Instantiate a Business Transformation Process for a Specific Organization.
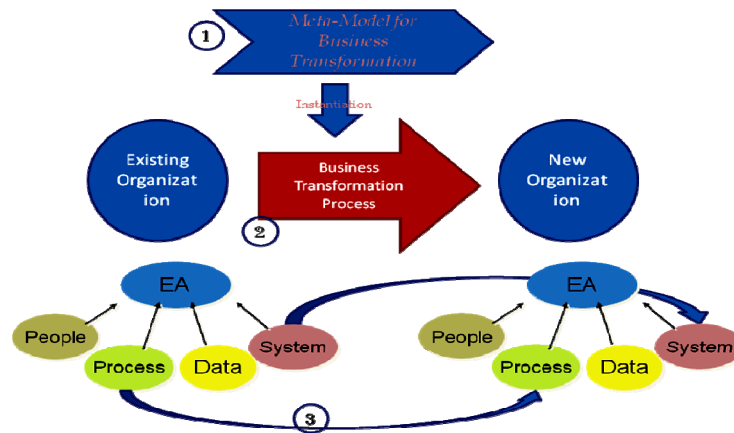
Figure 3: Correlating Business Transformation Process with Operational Business Processes undergoing transformation.

- Appoint roles to perform the activities.

Figure 3 shows three separate types of processes: (1) an abstract meta-model for business transformation (2) derivation and configuration of the specific transformation process and (3) the actual operational processes of the business undergoing transformation. Figure 3 further shows how an enterprise can be understood and modeled using an Enterprise Architecture (EA) framework. From the meta-model presented in earlier Figure 2 it can be seen that we need transformation processes to transform people, operational processes, Data and Systems. However, there can be many more elements depending on the organization and EA framework being used. Having identified these four core elements of an enterprise in this case, the derived transformation process can then transform each of these elements providing a holistic approach. Note that the operational business processes, in particular, need extensive modeling and walk-through for their successful transformation.

5 APPLYING BUSINESS TRANSFORMATION IN PRACTICE

The best way to demonstrate the application of business transformation in practice is through an example. Consider an organisation that is losing customers to its competition. The management of this medium sized organization has decided to counter this loss by introducing a CRM. The goal of this CRM is to improve responses to customer concerns and enhance customer experience. The business transformation meta-model provides many elements of which we select few that help in undertaking transformation of this SME. The process of business transformation is derived, configured and then enacted (shown in Figure 1 (B)) in order to transform the business. The following Table 1 summarizes the EA framework that could be used to understand and model the transformation of the aforementioned SME.

Table 1: Instantiation of the Core Elements of a Business Transformation meta-model for a Specific SME.

Elements of Transformation	Existing Organization	Transformed (Electronic) Organization	Expected Improved Outcomes and Comments
People: *Customers* *Senior management* *middle management* *Users*	Many employees, some repeating tasks. Customers have multiple touch points. Frustrated customers.	Single touch point for customers globally.	Greater value for the customer and reduced effort for employees. Training / up skilling is required
Process *Internal* *External*	Hierarchical; dispersed across many small departments. Unacceptable response time to customer issues.	Customer-centric processes. Improved response time. Enhanced customer experience.	Streamlining of processes through re-engineering and modeling results in efficient processes. Reduced wastage. Modeling and optimization required.
Data *Multi-media*	In multiple Access spreadsheets.	Converted into a single SQLserver database. Accessible on Internet and mobile devices.	Reduced duplication of data and hence less errors. Also, better control for the management. QA and testing required.
System *In-house* *Packages*	Multiple systems developed on 'as needed' basis.	Merging/Integration of systems into cohesive suite that is Internet-enabled. Able to support the people goal of single touch-point.	Less IT maintenance work and hence less costs. Reduction in energy budget. Improved performance. QA and testing required.

The elements of transformation listed in Table 1 are derived from a repository of transformation elements that are grouped as people, process, data and system (see Figure 3). This table further lists how the organization looks now and how it should look after transformation. This list, based on our transformation meta-model, provides starting point for a comprehensive business transformation and change management program. In case of the aforementioned example of an organization losing customers, following are the major steps in this transformation program:

- The organization is identified as a medium sized, localized organization
- People, process, data and system of the existing organization is listed (with the background of Zachman model, as described in Figure 3 of our model)
- Improving customer experience and thereby retain customers is the goal of the organization
- Integration of data and systems at technical level – using Web Services / Service Oriented Architecture (SOA)
- Up skilling / training of staff at people level.
- Customer management processes undergo re-engineering using BPM techniques (for example, see Ginige and Ginige(Ginige and Ginige, 2007))

- New business process architecture includes single point of contact (SPOC) functionality (see the comments in the Table 1).
- Transformation takes place for each element (described in each of the rows in Table 1) – e.g. people are trained, processes are re-engineering, data is converted (and tested) and systems are integrated keeping quality assurance in mind.

Eventually, this table expands into a comprehensive project plan that is based on the existing and future goals of the organization. Set of activities and tasks corresponding to the elements of Table 1 are incorporated in that project plan.

6 CONCLUSIONS AND FUTURE DIRECTIONS

This paper presents a framework to derive a holistic business transformation process. The goals of such framework is for the organisation to become agile and continuously be harmonious with the changes in the external business environment. The framework includes a meta-model which provides a suite or repository of various transformation elements made up of people, processes and deliverables. An EA framework was used to map the existing organization and its future state. We also described

the three different types of processes that are involved in transformation – the meta-model of a transformation process, the actual instance of the transformation process derived using the meta-model and the operational processes that are undergoing transformation. These elements are used in different context depending on the type and size of the organization and also its transformation goals. We plan to apply this transformation model to many other organizations in the region to further validate it.

The work presented in this paper will be extremely valuable to these organisations as research has shown many organisations do not clearly identify all aspects of the organisation which can be modelled through an EA that needs to be changed when undergoing a business transformation(Hol and Ginige, 2009). However, the practical application of this research will benefit by further action-research in the application of existing EA models (e.g. Zachman, TOGAF) from a transformation perspective.

REFERENCES

Arunathilaka, D. (2006) Applying Mobile Technologies to Banking Business Processes. *Handbook of Research in Mobile Business.* IGI Global, Hershey, PA, USA.

Arunatileka, D., Ghanbary, A. & Unhelkar B. (2008) Influence of Mobile Technologies on Global Business Processes in Global Organizations. IN RAISINGHANI, M. (Ed.) *Hand book of Research in Global Information Technology Management in the Digital Economy.* IDEAS Group Publication, Hershey, PA, USA.

Arunatileka, S. & Ginige, A. (2003a) Applying seven E's in eTransformation to the Manufacturing Sector. *cChallenges.* Bologna, Italy.

Arunatileka, S. & Ginige, A. (2003b) The Seven E'S in eTransformation - A Strategic eTransformation Model. *IADIS International Conference - e-Society 2003.* Lisbon, Portugal, International Association for Development of Information Society.

Deitel, H., Deitel, P. & Steinbuhler K. (2001) *e-Business and e-Commerce for Managers*, Prentice Hall, Upper Saddle River, New Jersey.

Ghanbary, A. & Unhelkar, B. (2007) Collaborative Business Process Engineering (CBPE) across Multiple Organisations in a Cluster. *IRMA Conference. IRMA 2007.* Vancouver, Canada, IRMA

Ginige, A. (2006) From eTransformation to eCollaboration: Issues and Solutions. In Yi-Chen, L. & Unhelkar, B. (Eds.) *2nd International Conference on Information Management and Business (IMB 2006).* Sydney, Australia.

Ginige, A. (2008) A Roadmap, Methodology and Technology Framework to support eTransformation. *4th International Conference on Information and Automation for Sustainability (ICIAfS 2008).* Colombo, Sri Lanka, IEEE Xplore.

Ginige, J. A. & Ginige, A. (2007) Challenges and Solutions in Process Automation in Tertiary Education Institutes: An Australian Case Study. *6th IEEE International Conference on Computer and Information Science (ICIS 2007).* Melbourne, Australia.

Hammer, M. & Champy, J. (2001) *Reengineering the Corporation – A manifesto for Business Revolution,* London, Nicholas Brealey.

Hol, A. & Ginige A. (2008) Dimensions of eTransformation. *4th International Conference on Information and Automation for Sustainability (ICIAfS 2008).* Colombo, Sri Lanka, IEEE Xplore.

Hol, A. & Ginige, A. (2009) eTransformation Journey for SMEs. *International Joint Conference on e-Business and Telecommunications (ICE-B 2009).* Milan, Italy, INSCITT.

Kalakota, R. & Robinson, M. (2001) *e-Business 2.0: Roadmap for Success (2nd Edition)* Addison-Wesley, USA.

Lan, Y. & Unhelkar, B. (2005) *Global Enterprise Transitions*, IGI Global, Hershey, PA, USA.

The_Open_Group (2010) The Open Group Architectural Framework (TOGAF)

Unhelkar, B. (2003) *Process Quality Assurance of UML based projects*, Addison-Wesley, Boston.

Unhelkar, B. (2009) Mobile Business: Upcoming Opportunities, Challenges, and Transition Approaches. *Cutter Benchmark Review,* 9**,** 13-31.

Unhelkar, B. (2010) Managing Mobile Business. IN SIMON, P. (Ed.) *The Next Wave of Technologies: Opportunity From Chaos.* John Wiley and Sons.

Unhelkar, B., Ghanbary, A. & Younessi, H. (2009) *Collaborative Business Process Engineering and Global Organizations: Frameworks for Service Integration*, IGI Global, Hershey, PA, USA.

Zachman, J. & Holcman, S. (2010) Zachman Framework

SHORT PAPERS

ASSESSING BUSINESS TRANSACTION STANDARDS AND THEIR ADOPTION
A Cross Case Analysis Between the SETU and Vektis Standards

Wouter Berends

Department of Business Information Systems, TNO ICT, Enschede, The Netherlands
Wouter.berends@tno.nl

Erwin Folmer

Department of Information Systems and Change Mangement, University of Twente, Enschede, The Netherlands
Department of Business Information Systems, TNO ICT, Enschede, The Netherlands
erwin.folmer@tno.nl, e.j.a.folmer@utwente.nl

Abstract: Nowadays businesses increasingly want to be interoperable so that they can collaborate with other organizations. Interoperability can be achieved through the use of business transaction standards, by which the organizations that use the standards collectively form a value added network. However the effectivity of these standards is largely dependant on the number of organizations that have adopted it, and thus it is very important that the standard conforms to the conditions that organizations have towards adopting these standards. Building on recent literature describing technical standards (van de Kaa, 2009), we have constructed a model through which standard aspects can be compared with the adoption conditions that organizations have. Subsequently cross case analysis methods were used to identify important aspects that influence the adoption of business transaction standards, as well as the identification of methods by which the aspects can be adapted by an Standard Development Organization (SDO) so that higher standard adoption is achieved. This evaluation can give managers and SDO's a higher understanding on standards itself and the domain it is supposed to function in. The cases demonstrated that early involvement of organizations with high market powers (preferably through a federation that represents these organizations) is important for adoption whereby the development and maintenance of the standard should preferably be funded by those organizations that have most to gain from broad standard adoption. Furthermore open characteristics, modularity and efficient business processes are perceived imperative for the adoption of business transaction standards.

1 INTRODUCTION

Business transaction standards are used by Inter - Organizational information Systems (IOS) to increase the level of interoperability amongst collaborating organizations. Collaborative business (C-business) describes the interlinked collaboration of all participants in a value added network (Scheer, 2003) through which organizations agree to work together as a method for achieving their common goals. When collaborative business is strived after, the systems of every party involved must be able to interoperate. This interoperability can be achieved through a standard that is recognized by all involved parties. The advantages that can be derived from standard use is that documents and messages are delivered to the target organization in such manners that they can automatically be processed through their business processes, without necessarily owning and controlling the asset. "Through this unique combination of resources advantage over competing firms can be established. These unique interfirm linkages may be a source of relational rents and competitive advantage" (Dyer, 1998). Because standards enable interfirm linkages the relational view as described by Dyer (1998) is applied who has identified several advantages (relation-specific assets, interfirm knowledge sharing routines, complementary resources / capabilities and effective

151

governance) resulting from interconnecting resources. Existing standard research suggest that standard adoption is dependant on the aspects that that standard have (van de Kaa, 2009), however most research describe technical standards (like the USB and IPv6 standards). Research describing business transaction standards is scarce (Folmer, 2009). Furthermore because of political, organizational, and economic complexities (Backhouse, 2003) business transaction standard development differ from technical standards and thus indicates that the aspects form technical standards are not necessarily the same as business transaction standards. The goal of this paper was to identify those aspects that are of importance for the adoption of business transaction standards, and search for methods through which these aspects can be adapted by an Standard Development Organization (SDO) so that higher standard adoption is achieved. In contrast to van de Kaa (2009) who developed a framework showing technical standards aspects through which managers can make a well informed choice on which standard to use. We argue that an SDO can also use the aspects to alter the standard in order to achieve higher adoption degrees and subsequently achieve higher interoperability. This was achieved by constructing an adoption model consisting of adoption factors found in existing literature. The model was validated by using cross case evaluation techniques, which was done by first constructing case descriptions of the two business transaction standards. Second using the case descriptions cross case analysis has been conducted that show similarities and differences between the two cases. And third using the cases a determination was made which aspects are of importance for business transaction standards as well as making adoption hypotheses.

2 STANDARD ADOPTION CONDITIONS

Organizations can have multiple factors that encourage or restrict them for adopting business transaction standards (i.e. adoption conditions), for the SDO it is important to know which adoption conditions are present so that the standard aspects can be altered in such manners that the standard complies to those conditions. In order to ascertain these conditions multiple organizational viewpoints were adopted whereby the organizational viewpoints are those organizational types involved in standards.

Krechmer (2009) identified end users, implementation organizations and SDO's as those organizational types that have different viewpoints towards adoption conditions. The end user organization will benefit from the business transaction standard because the application of IT and complementary organizational resources can improve business processes / enable new ones and may ultimately impact organizational performance. Zhu (2006) identified two business drivers that are of key influence to the adoption of end user organizations: Network effects and switching costs. The implementation organization is the organization that makes software (in-house or outhouse) that is compliant to the standard which is developed by the SDO. The implementation organization ensures that the end user is able to use the software / standard. Since our goal is to assess the standard that is used by that specific software solution, a software supplier (portability) and maintainer (maintainability) view has been used and not a software developer view. The SDO usually is a non-profit organization that is comprised with members that have common goals, thus the goal of an SDO is the development and maintenance of a standard that service the common interests of it's members. Zhu (Zhu, 2006) describes that for an SDO one measure is key for adoption and diffusion which is deployment. The deployment of the standard can subsequently be judged based upon three measures: Volume, Diversity and Breadth.

3 STANDARD ASPECTS

Following the assumption that the aspects identified for technical standards are not the same as the aspects of influence for business transactions standards, we have identified the aspects that are of importance when dealing with business transaction standards. In order to establish a complete aspect overview that influence the adoption of business transaction standards, articles that discuss business transaction standard aspects have been examined. Starting point was the model by van de Kaa (2009) who has made a similar model describing factors for standard dominance specific for network (technical) standards. Additional aspects were identified through other literature sources. These articles were published in the top 25 CS/IS journals (Mylonopoulos, 2001) and the top 25 International Business Journals (DuBois, 2000). Second articles discussing business transaction standard case studies were used to examine which aspects were mentioned

relevant to adoption. This resulted in twenty six aspects that were subsequently subdivided into six distinct categories derived from Melville (2004), the aspects are shown in Table 2. The importance of the Melville model is that it encompasses every domain (e.g. Focal firm, Competitive Environment and Macro environment) that Information Technology have influence upon. By embracing these categorizations it was possible to conclude that all aspects have been identified furthermore these categorizations ensured that generalizations and aspects pattern recognition can be used when applying case evaluation methods.

4 CROSS CASE EVALUATION

In order to ascertain which standard aspects can best be influenced by SDO's so that higher adoption degrees are achieved, cross case analysis methods have been used. These cross case analysis methods have yielded qualitative results through which an in depth understanding was ascertained regarding the standard aspects that influence the organizational adoption conditions. The standards used for this analysis were the Vektis (http://www.vektis.nl) standard and the SETU (http://www.setu.nl) standard. Whereby the goal of this cross case analysis was to identify those aspects that are of importance of standard adoption, and to determine how these aspects relate to the adoption conditions. For both the SETU case and Vektis case a preliminary case description has been constructed using literature which has been provided by the SETU and Vektis organizations itself, internet (site) and Dutch articles. These sources provided a basis from which an understanding was formed regarding the dimensions that exist when using the standards. Because the general goal of this paper is to form hypotheses and to search for adoption patterns amongst standard aspects, the second data collection phase was conducted using separate data collection methods (Eisenhardt, 1989).

- The data used for the Vektis case study has been formed through a questionnaire that was sent to Vektis KEI members. These members represent both end user and implementation organizations and are involved with the standard build (43% response rate). This questionnaire consisted of open questions derived from literature, aspect categories and the adoption conditions.

- The data used for the SETU case study was formed through interviews with TNO employees, who are involved with the building and maintenance of the standard. As well as reviewing documentation provided by TNO.

The following two paragraphs will give short descriptions of the two standards.

4.1 Vektis Case

Vektis is an organization that is funded by insurance organizations and ensures that appropriate information is available for health providers so they can perform their tasks in conjunction with their business (chain) partners. One of their activities is to provide standards for the declaration processes in the Dutch healthcare domain, by which standard users are insurers, health offices and health providers. The prime task of the standard is to provide mechanisms that show whether individuals are insured and to provide mechanisms that enable electronic health declarations. Whenever patients have used services provided by health providers a billing process is initiated towards the insurer. These declarations are transmitted electronically towards the insurance companies which in turn provide payment towards the health providers. The transmissions go through the VECOZO portal that ensures that the overall communication is conducted in a secure and safe manner. Vektis is also placed in the Dutch environment as an independent organization, by which it's partners, the Dutch health insurers, provide it's financial backing.

4.2 SETU Case

The SETU (Foundation for electronic transactions in the staffing industry) was founded by the Dutch federation for the staffing industry (ABU). The SETU standards serve the process between staffing organizations and organizations that acquire personnel through the mediation of the staffing organizations. The general goal of the standard is to facilitate electronic transactions between the organizations within the Dutch staffing industry, to standardize the business process for compatibility reasons and to ensure continuity of the developed standards. The SETU standards are in existence since the beginning of 2007, since then the standards handle approximately 10 percent of all timecards.

Table 1: Field case results.

	Vektis	Setu
IT Resources	The standards are relatively open, whereby the insurance organizations pay contributions. Standards modules are built for specific occupational groups, resulting in high recognition, and increases the adoption amongst those groups. Furthermore the standards are built using ASCII and it is not compatible with other standards, however this has no influence on the adoption. The standards are also considered easy to use by which Vektis actively encourages the testing of the standards which is considered important for it gives a sense of confidence.	Although the standards are open the standards are not as widely used that software vendors have solutions ready "on the shelf" which are SETU compliant, this stresses the switching costs and portability. However many organizations / software solutions are already compliant to HR-XML which ensures that lower switching costs and portability is mediated and also linkages can be made to international organizations, The standards are built out of four complementary standards by which organizations can choose to implement one or multiple standards.
Compl. Org. Resources	Broad adoption of the standards can mainly be attributed to cost savings organizations within the healthcare domain can make significant cost savings through the standards. The fact that Vektis and Vecozo is funded by insurance organizations does not affect the adoption of the standards. However the presence of the KEI within Vektis encourages good will and reduces misinterpretations and thus has positive effects on adoption.	The development of the standards is conducted in such a manner that organizations can participate in the development. This creates good will and reduces misinterpretations amongst industry players. Furthermore organizations within the staffing domain can make significant cost savings through the usage of the standards (economies of scale do apply).
Business Processes	The standard ensures that the declaration process is conducted on an unambiguous manner which makes the adoption of the standard attractive because it can be used to enhance business processes. Furthermore insurance organizations also offer to collect the money that is not covered by the insurance, which creates good will and makes the declaration process simpler.	The standards ensure that no paper documents (e.g. hour specifications) will have to be sent towards other organizations that have converted to the standards. This will yield burden reliefs through increased processes which can attract organizations to adopt the standards.
Industry Characteristics	The healthcare declaration processes is highly regulated within the market which results in high adoption. By which health insurance organizations do not compete by means of an effective communication process between chain partners, they do compete through lower insurance fees and by having superior customer contacts etc. This means that the health insurance organizations can collectively endorse / develop the standards.	When adopting the SETU standards staffing customer organizations have the possibility of reaching more staffing companies, and thus giving them more flexibility. This can have positive effects on the adoption of the standards. However the staffing industry is one where there is lots of competition (i.e. the market is not consolidated) and thus organizations can choose not to convert in order to be flexible to market changes.
Trading Partners	Insurance organizations collectively have such powers that they can obligate (by means of contractual agreements) the usage of the standards, and thus health providers will have to implement the standards. The high adoption of the standards has ensured that the programs (that are compliant to the standards) are relatively cheap and thus lowers the switching costs.	At this point in time the installed base of the standard is not of such levels that other organizations automatically feel the need to adopt, there are several important organizations (e.g. Randstad, Adecco) within the market that are actively working for higher adoption of the standards, however they cannot make the standard use compulsory towards their partners.
Macro Env.	The standards are not obligated by law, however the healthcare domain is strongly regulated which is one of the reasons why the insurance organizations have such powers in the market. Furthermore respondent stated that when the government would have obligated standard use adoption would have gone quicker.	The SETU standards have been added to the comply-or-explain regime of the Dutch government which pressures (semi) governmental organizations to convert to the SETU standards, furthermore expectations are that this regime will encourage other organizations to follow.

Recently the standard has been added to a Dutch government "Comply-or-Explain" list that includes open standards that are mandatory to be used for each (semi) government organisation for achieving interoperability (Min_EZ, 2007).

4.3 Cross Case Vektis and SETU

Using the data collected from the SETU and Vektis standards, a cross case analysis has been conducted whereby aspects that are important for the adoption of business transaction standards have been identified. Table 1 shows the similarities and differences between the two cases. These differences between the two cases can mainly be related to the competitive environment and the macro environment the main differences are:

- The staff lending domain is not consolidated resulting on more emphasis on the IT resource dimension e.g. making the standard more compatible with other standards.

- Legislation ensures that health insurance organizations have high domain powers, as a result the Vektis standards are purely built to serve the declaration process towards the

Table 2: Aspects versus organizational conditions.

Standard aspects	End user organization		Implementation organization		SDO		
Focal Firm - IT Resources	1	2	3	4	5	6	7
1.1 Technological superiority	L	L		L			
1.2 Compatibility of standard	L & F	F	L & F	L & F	L		L
1.3 Open standard	L & F	L & F	L	F	L & F		
1.4 Complete set of functionalities	L & F		F	L& F	L& F	F	L
1.5 Customization capabilities	L	L		L		F	
1.6 Ease of use				L & F	L & F	L	
1.7 Deployment strategy	L			L & F	L		L
Focal Firm - Complementary Organizational Resources							
2.1 Financial strength	L & F				L& F		L
2.2 Complementary goods and reputation	F				F		
2.4 Participation in standard consortia	L & F				L & F		L
2.5 Pricing strategy that attracts customers		L					
2.6 Customer expectations					L		
2.8 Management support	L			L			
Focal Firm - Business processes (performance)							
3.1 Process management is increased	L & F				L & F	F	L
3.2 Business model is extended	L & F				L		L
Competitive Environment - Industry characteristics							
4.1 Vertical integration	L & F				L		L & F
4.3 Market is ready	L & F				L & F		
Competitive Environment - Trading partners							
5.1 Network externalities	L	F			L & F		L
5.3 Big Fish	L & F				L & F	F	
5.4 Stakeholders in standard build	L & F						L
Macro Environment							
6.1 Legislation that encourages standard usage	F				L & F	F	F

(1 = Network effects, 2 = Switch costs, 3 = Maintainability, 4 = Portability, 5 = Volume,6 = Diversity, 7 = Breadth)
(F= Aspect found in field case (this paper), L = Aspect described in literature cases (earlier work)

- insurance organizations, making the characteristics of the standard less important.

Table 2 shows the aspects that were deemed important (after assessing the SETU and Vektis cross case analysis) per organizational adoption condition.

5 CONCLUSIONS

The goal of this paper was to identify those aspects that are of importance for the adoption of business transaction standards, and search for methods through which the aspects can be adapted by an SDO so that higher standard adoption is achieved. During the cross case analysis it has become apparent that the main reason for organizations to adopt business transaction standards is highly dependant on the powers that exist in the specific domain. When there are large power differences amongst organizations, then the SDO should focus on those aspects that reside in the competitive environment category (table 2). Whenever there are

low differences then the focus should be towards the IT resource category (table 2). In all cases it is important that the standard is developed by a federation that represents organizational (types) residing in the domain, as well as basing the standard on open characteristics. This will give organizations a sense of ownership and makes the standard free to use, which is good for adoption. Furthermore the standard build should be modular whereby the modules should be based on best practice business processes, enabling organizations to only adopt those modules that are of importance for them. SDO organizations should produce standards in such manners that it fulfils the conditions that organizations have set and thus increase the chance that the standard will be adopted. Following the conclusions made by (van de Kaa, 2009) who has made a similar model describing important aspects for technical standards, the model described in table 2 can be used by managers in order to ascertain a deeper insight into those aspects that are of influence for their organization and subsequently make decisions which standard their organization should support.

REFERENCES

Backhouse J., Hsu C., McDonnell A. (2003) "Toward public-key infrastructure interoperability" Communications of the ACM 46(6): 98-100.

Benner, M.J. and M.L. Tushman (2003) "Exploitation, exploration, and process management: The productivity dilemma revisited", Academy of Management Review 28(2): 238-256.

DuBois, F. L. and D. Reeb (2000). "Ranking the international business journals." Journal of International Business Studies 31(4): 689-704.

Dyer J.H., Singh H. (1998) "The relational view: Cooperative strategy and sources of interorganizational competitive advantage" Academy of Management Review 23(4): 660-679.

Eisenhardt K.M., 1989. "Building Theories From Case Study Research" Academy of Management, The Academy of Management Review 14(4): 532-550.

Folmer E.J.A., Berends W.B., 2009. "Top IS Research on Quality of Transaction Standards" The 6th International Conference on Standardization and Innovation in Information Technology.

Kaa G. van de, 2009 "Standards Battles for Complex Systems - Empirical Research on the Home Network" Reference number ERIM: EPS-2009-166-ORG ISBN 978-90-5892-205-2.

Krechmer, K. (2009) "Open standards: A call for change" IEEE Communications Magazine 47(5): 88-94.

Lu X.H., Huang L.H., Heng M.S.H. (2006) "Critical success factors of inter-organizational information systems - A case study of Cisco and Xiao Tong in China" Information and Management 43(3) 395-408.

Melville N. (2004), REVIEW: Information Technology and Organizational Performance: An Integrative Model of IT Business Value, MIS Quarterly Vol. 28 No. 2 pp283-322.

Ministerie van Economische Zaken, 2007 Nederland Open in Verbinding Een actieplan voor het gebruik van Open Standaarden en Open Source Software bij de (semi-) publieke sector, retrieved 08-2009 from, http://appz.ez.nl/publicaties/pdfs/07ET14l.pdf

Mylonopoulos, N. A. and V. Theoharakis (2001). "Global perceptions of IS journals - Where is the best IS research published?" Communications of the Acm 44(9): 29-33.

Scheer, A.-W., Grieble O., Zang, S. (2003) "Collaborative Business Management" In: Kersten, W. (ed.): E-Collaboration - Prozessoptimierung in der Wertschöpfungskette. Deutscher Universitäts- Verlag, Wiesbaden, 30 et seq: 29-57.

Zhu K.,Kraemer K.L.,Gurbaxani V.,Xu S.X. (2006) "Migration to open-standard interorganizational systems: Network effects, switching costs, and path dependency" MIS Quarterly: Management Information Systems 30(SPEC.ISS.): 515-539.

DESIGNING ENTERPRISE ARCHITECTURES BASED ON SYSTEMS THEORETIC STABILITY

Philip Huysmans, David Bellens, Dieter Van Nuffel and Kris Ven

Department of Management Information Systems, University of Antwerp, Prinsstraat 13, B-2000 Antwerp, Belgium
{philip.huysmans, david.bellens, dieter.vannuffel, kris.ven}@ua.ac.be

Abstract: Contemporary organizations are operating in increasingly volatile environments. Hence, organizations must be agile in order to be able to quickly adapt to changes in its environment. Given the increasing complexity of organizations, it has been argued that organizations should be purposefully designed. Enterprise architecture frameworks provide guidance for the design of organizational structures. Unfortunately, current enterprise architecture frameworks have a descriptive, rather than a prescriptive nature and do not seem to have a strong theoretical foundation. In this paper, we explore the feasibility of extending the prescriptive design principles of the Normalized Systems theory to the field of enterprise architecture. Our results show that such approach is feasible and illustrate how the systems theoretic concept of stability can be used on the organizational level.

1 INTRODUCTION

Contemporary organizations are operating in increasingly volatile environments. Hence, organizations must be agile in order to be able to quickly adapt to changes in their environment. This may be a complex process, since a change to one organizational unit may affect other units. Given the increasing complexity of organizations, it has therefore been argued that organizations should be purposefully *designed* in order to exhibit true agility (Hoogervorst, 2009). Enterprise architecture frameworks support the design of the organizational structure, its business processes and information systems through a coherent set of principles, methods and models (Bernus et al., 2003). Unfortunately, current enterprise architecture frameworks have a descriptive, rather than a prescriptive nature. In order to purposefully design organizations, prescriptive principles are needed.

In software engineering literature, the Normalized Systems approach has recently been proposed to provide such deterministic design principles for the modular structure of software. The Normalized Systems approach is based on the systems theoretic concept of stability to ensure the evolvability of information systems. It argues that the main obstacle to evolvability is the existence of *combinatorial effects*. Combinatorial effects occur when the effort to apply a specific change increases as the system grows. This is a result of Lehman's law, which states: "*As an evolving pro-*

gram is continually changed, its complexity, reflecting deteriorating structure, increases unless work is done to maintain or reduce it." (Lehman, 1980). The Normalized Systems approach eliminates these combinatorial effects by defining clear and deterministic design principles. Adhering to these principles therefore results in software systems that exhibit stability.

In this paper, we extend the Normalized Systems approach to the domain of enterprise architecture. The issue of combinatorial effects has not previously been explored in enterprise architecture frameworks, but seems relevant. Also, applying systems theory to the construction of organizations would advance the emerging field of enterprise engineering (Liles et al., 1995). We therefore explore the feasibility of designing a method for the construction of enterprise architectures that exhibit systems theoretic stability by eliminating combinatorial effects. Such a method would provide a more deterministic way of designing agile organizations. This method is currently being developed by using the design science methodology.

2 ENTERPRISE ARCHITECTURE

When market threats, opportunities or changes arise, the organization as a whole has to adapt. In order to be able to comprehend and manage the complexity of modern organizations, enterprise architecture frameworks have been introduced. Despite the common

goal of enterprise architectures, many different frameworks are available. Various authors (e.g., (Leist and Zellner, 2006)) have compared these frameworks and identified differences and similarities. According to Leist and Zellner, who evaluated enterprise architecture frameworks with regard to the requirements of method engineering, no framework exists which provides all necessary elements to constitute a complete method (Leist and Zellner, 2006). Should an enterprise architect require the use of all elements, several (complementary) frameworks can be used concurrently, or a particular framework can be extended with missing elements. However, by combining or extending existing frameworks, the issue of *integration* between the models in the framework becomes even more complex. While most frameworks reduce the inherent complexity of an organization by offering separate views, it is not always clear how these views relate to or affect each other. The integration between the conceptual models should facilitate the translation of a single change in the outside world to all the different aspects of the organization.

However, if a change in a certain model affects other models it is combined with, a *combinatorial effect* occurs. While originally used in the Normalized Systems approach to describe evolvability in software, combinatorial effects also seem to affect evolvability on the enterprise architecture level. Analogously with combinatorial effects on the software level, this implies that organizations would become less evolvable as they grow. While the issue of integration has been acknowledged by other authors (e.g., (Lankhorst, 2005)), it has, to our knowledge, not yet been studied based on system theoretic concepts such as stability. By applying the design principles from Normalized Systems to enterprise architecture, we attempt to introduce these concepts in this field. In this paper, we elaborate on the construction of the *core diagram*. The core diagram is a model which provides an overview of the organizational scope which will be designed (Ross et al., 2006). Moreover, the core diagram aids understandability and communication of an enterprise architecture framework.

3 THEORETICAL FOUNDATION

3.1 Enterprise Ontology

Enterprise Ontology views the organization as a social system (Dietz, 2006). Therefore, it is well suited to describe the interaction between an organization and its environment. Enterprise Ontology assumes that communication between human actors is a neces-

sary and sufficient basis for a theory of organizations (Dietz, 2006). This is based on the language action perspective and Habermas theory of communicative action. The strong theoretical foundation ensures a consistent modelling methodology. Clear guidelines are provided to create abstract models. Since only the *ontological* acts are represented in the models, the same model will be created for organizations who perform the same function, but operate differently. For example, consider the BPR case at Ford (Hammer, 1990). The ontological model of the processes of the situation before and after reengineering are identical. Because of the focus on the essential business processes, Enterprise Ontology models can be very concise. Therefore, they provide a good overview of a broad enterprise scope, and are well suited as an enterprise architecture core model.

The transaction pattern describes the coordination necessary to produce a certain result. This result is represented by a *production fact*. There are always two actors involved in a transaction: the *initiator* actor who wants to achieve the fact, and the *executor* actor who performs the necessary actions to create the fact. Delivering a product, performing a service or subscribing to an insurance are examples of production facts which could be created by completing a transaction. The high-level structure of the transaction pattern consists of three phases. In the order phase, the actors negotiate the subject of the transaction. In the execute phase, the subject of the transaction is brought about. In the result phase, the result of the transaction is presented and accepted. In different versions of the transaction pattern, different ontological process steps are identified in the three phases. These steps are called *coordination acts*. The successful completion of an act results in a *coordination fact*.

The basic transaction pattern consists of the five standard acts which occur in a successful scenario (i.e., request, promise, execute, state and accept) (Dietz, 2006, p. 90). Consider a transaction in the case of a simple product delivery process. In the order phase, the customer *requests* the product. Once this request is adequately specified, the request coordination fact is created. The supplier then *promises* to deliver the product according to the agreed terms. This creates the promise coordination fact. In the execute phase, the executor actually performs the the execute act, resulting in the production fact. In our example, this is the actual delivery (i.e., "Product X has been delivered"). In the result phase, the supplier *states* that the delivery has been completed. If the customer is satisfied with the delivery, he will accept the delivery in the *accept* process step. Once the accept coordina-

tion fact is created, the transaction is considered to be completed.

The *standard* transaction pattern is the basic transaction pattern, augmented with the scenario in which the actors dissent (Dietz, 2006, p. 93). In the order-phase, the executor actor can decline the incoming request of the initiator actor. The initiator then has to decide whether he resubmits his request, or quits the transaction. In our example, the supplier could decline the delivery of a product which does not belong to his catalogue. The customer would need to select another product, or quit the transaction and search another supplier. The execute-phase is identical to the execute-phase in the basic transaction pattern. In the result-phase, the initiator actor can reject the stated production fact instead of accepting it. The executor then has to decide whether he wants to repeat the execution act and make the statement again, or stop the transaction.

3.2 Normalized Systems

The basic assumption of the Normalized Systems approach is that information systems should be able to evolve over time, and should be designed to accommodate change. To genuinely design information systems accommodating change, they should exhibit *stability* towards requirements changes. In systems theory, stability refers to the fact that bounded input to a function results in bounded output values, even as $t \to \infty$. When applied to information systems, this implies that no change propagation effects should be present within the system; meaning that a specific change to an information system should require the same effort, irrespective of the information system's size or the point in time when being applied. *Combinatorial effects* occur when changes require increasing effort as the system grows. They need to be avoided in stable systems. Normalized Systems are therefore defined as information systems exhibiting stability with respect to a defined set of changes (Mannaert and Verelst, 2009), and are as such defying Lehman's law of increasing complexity (Lehman, 1980) and avoiding the occurrence of combinatorial effects.

The Normalized Systems approach proposes a set of four *design principles* that act as design rules to identify and circumvent most combinatorial effects (Mannaert and Verelst, 2009). The first principle, *separation of concerns*, implies that every change driver or concern should be separated from other concerns. This theorem allows for the isolation of the impact of each change driver. The second principle, *data version transparency*, implies that data should be com-

municated in version transparent ways between components. This requires that this data can be changed (e.g., additional data can be sent between components), without having an impact on the components and their interfaces. The third principle, *action version transparency*, implies that a component can be upgraded without impacting the calling components. This principle can be accomplished by appropriate and systematic use of, for example, polymorphism or a facade pattern. The fourth principle, *separation of states*, implies that actions or steps in a workflow should be separated from each other in time by keeping state after every action or step. This suggests an asynchronous and stateful way of calling other components.

The design principles show that software constructs, such as functions and classes, by themselves offer no mechanisms to accommodate anticipated changes in a stable manner. The Normalized Systems approach therefore proposes to encapsulate software constructs in a set of five higher-level software elements. These elements are modular structures that adhere to these design principles, in order to provide the required stability with respect to the anticipated changes (Mannaert and Verelst, 2009). From the second and third principle it can straightforwardly be deduced that the basic software constructs, i.e., data and actions, have to be encapsulated in their designated construct. As such, a *data element* represents an encapsulated data construct with its get- and set-methods to provide access to their information in a data version transparent way. So-called cross-cutting concerns, for instance access control and persistency, should be added to the element in separate constructs. The second element, *action element*, contains a core action representing one and only one functional task. Four different implementations of an action element can be distinguished: *standard* actions, *manual* actions, *bridge* actions and *external* actions. In a standard action, the actual task is programmed in the action element and performed by the same information system. In a manual action, a human act is required to fulfil the task. The user then has to set the state of the life cycle data element through a user interface, after the completion of the task. A process step can also require more complex behaviour. A single task in a workflow can be required to take care of other aspects, which are not the concern of that particular flow. Therefore, a separate workflow will be created to handle these concerns. Bridge actions create these other data elements going through their designated flow. When an existing, external application is already in use to perform the required task, the action element would be implemented as an external action.

These actions call other information systems and set their end state depending on the external systems' reported answer. Based upon the first and fourth principle, workflow has to be separated from other action elements. These action elements must be isolated by intermediate states, and information systems have to react to states. To enable these prerequisites, three additional elements are identified. A third element is thus a *workflow element* containing the sequence in which a number of action elements should be executed in order to fulfill a flow. A consequence of the stateful workflow elements is that state is required for every instance of use of an action element, and that the state therefore needs to be linked to or be part of the instance of the data element serving as argument. A *trigger element* is a fourth one controlling the states (both regular and error states) and checking whether an action element has to be triggered. Finally, the *connector element* ensures that external systems can interact with data elements without allowing an action element to be called in a stateless way.

4 ARTEFACT CONSTRUCTION

In this section, we outline the construction of a core diagram which is based on Enterprise Ontology models and expressed in Normalized Systems constructs. Since Enterprise Ontology models are implementation-independent, we can base our method on these models to implement the needed organizational aspects in the transactions. The resulting artefact is called a *Normalized Systems Business Transaction (NSBT)*. In order to illustrate the different steps, we use a mail order example. In this example, different implementations of ontological process steps are available. For example, consider the request of the order. Instead of using a standard mail form, the company can offer the customer the possibility to place the order on a website. We introduce these variations to illustrate the evolvability of the NSBT with regard to changes in implementation technology. While this changes the implementation, no changes are made to the *essential* Enterprise Ontology models.

In the mail order example, an Enterprise Ontology transaction would result in a production fact "*an order has been delivered*". In Normalized Systems, this transaction pattern is represented by a flow element. A flow element is driven by precisely one data element, the life cycle data element. In order to define a Normalized Systems flow, we thus need a T01 data element. The completion of the different acts in the transaction process is represented by the creation of ontological facts. In Normalized Systems, these facts are represented by the states which occur in the flow element, being the life cycle states of the corresponding data element. To reach these states, a state transition is required. A state transition is realized by an action element. The successful completion of that action element results in the defined life cycle state. In order to define the control flow of the process, we therefore need to specify the *trigger states*, *state transitions* and *transaction actions*. Regarding the request coordination fact, this implies that the T01 flow element, and thus also the corresponding T01 data element, should reach the state *Requested*. This means that upon initiation of a T01 transaction, a new T01 data element is instantiated, resulting in the life cycle state *Initial*. The genuine act of requesting is encapsulated in the action element Request. In the mail order example, the Request action element would contain the task which ensures that the order request is fully defined by the customer. In our automated version of the example, this functionality is offered by the website. However, if the traditional mail order request form needs to be supported as well, a second Request action element could be created. This action element would be implemented as a manual action. When a retail company employee receives the order request form, the *Requested* state will be set manually. The remainder of the transaction will be handled identically, regardless of the implementation method of the request. The *concerns* of creating the data element and handling the request are thus separated as they can clearly evolve independently from each other.

While all state transitions are defined as action elements, their different nature can mean that they need to be implemented differently. Consider the notification of the initiator actor in the promise process step. If this notification requires a human action, the Promise action element would be implemented as a manual action. For example, it could be that the order request needs to be approved by an employee of the retail company. However, the promise process step can also require more complex behaviour. When for example the requested product first needs to be reserved in the warehouse, the Promise action element would be implemented as a bridge action triggering a flow element on another data element, e.g., a Reservation element. When the retail company already has an existing application in use to perform these reservations, the Promise action element would be implemented as an external action.

However, the transaction process does not always follow the successful scenario. In the scenario in which the actors can dissent, additional coordination acts need to be added. When translating these acts to
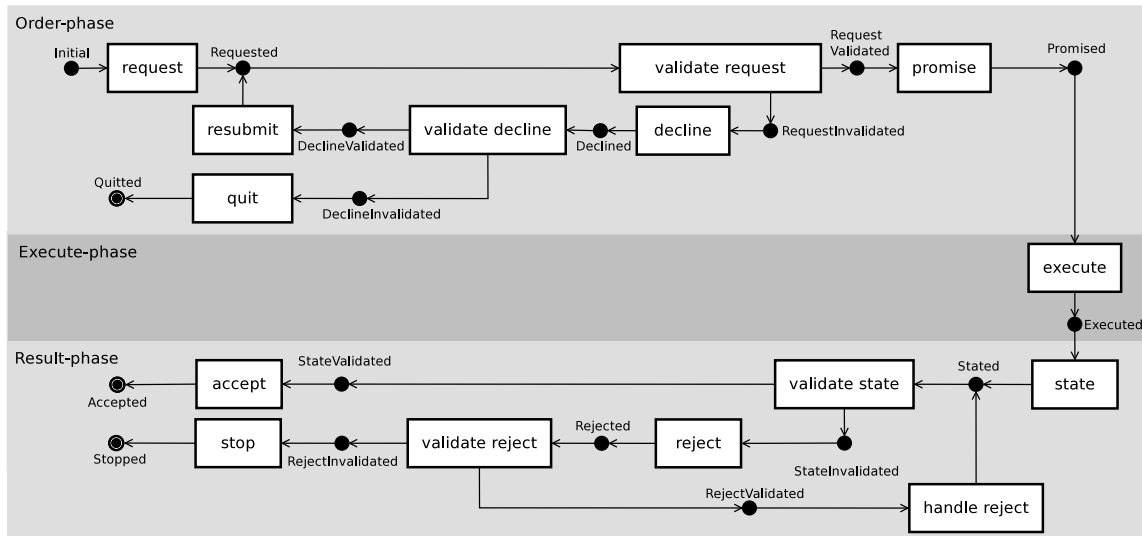
Figure 1: Graphical representation of the standard transaction pattern flow.

Normalized Systems primitives, some additional actions and states have to be included due to the Normalized Systems theorems. The resulting Normalized Systems flow element is graphically represented in Figure 1. Based on *separation of concerns*, the decision of the executor actor to promise or decline the request needs to be separated from the actual coordination act (i.e., the communication of the decision). The decision logic to promise or decline can change independently from the communication method, as shown by the various implementations of the `Promise` action element described above. Since the communication method can also change independently of the decision logic, these two actions should not be combined in one action element. Doing so would introduce a combinatorial effect. Therefore, we introduce an additional action element `ValidateRequest`. In the case where the executor decides to handle the request, the state *RequestValidated* is set. Otherwise, the state *RequestInvalidated* is set. The actual `Promise` action element then contains the actual communication of the decision. In our example, the `ValidateRequest` action element can contain the logic to check whether the retail company can deliver the order, e.g., whether the product is in stock. If the request is declined, the initiator actor needs to decide whether or not to resubmit the request. This decision logic is again separated from the other actions by encapsulating the decision logic in an action element `ValidateDecline`. If the initiator decides to resubmit, the state is set to *DeclineValidated*. The `Resubmit` action element then allows the initiator actor to possibly change the request and to resubmit it which will again result in the state *Requested*. If the product from the original order request is not available, the customer thus has the

option to adapt his order and resubmit it. If the initiator decides to abort the transaction, the state is set to *DeclineInvalidated*, which triggers the `Quit` action element to reach the end state *Quitted*. Analogously, the initiator actor has to decide whether he accepts the stated production fact in the result phase of the transaction. We therefore introduce the `ValidateState` action element, which results in the *StateValidated* state in case of a successful acceptance, or in the *StateInvalidated* state in case of an unsuccessful one. The *StateValidated* state triggers the `Accept` action element, which contains the actual accept coordination act. In case the initiator does not accept the state coordination fact, the workflow is brought to the *Rejected* state through the `Reject` action element. The separation of concerns theorem forces us again to separate the action element containing the decision logic (i.e., the `ValidateState` action element) from the action element containing the communication method (i.e., the `Accept` and `Reject` action elements). In our example, it is possible that the customer is not satisfied with the delivered products. This decision can be implemented as a manual action element. The user would manually check the delivery, and indicate whether he accepts it. However, in the context of a B2B transaction, it could be that an automated quality control system is in place. In that case, an external action element would be used. These different action elements could bring the state of the transaction workflow in the *StateInvalidated* state. The `Reject` action element, which communicates the decision, can again be implemented using the different types of action elements. The decision whether to handle the reject is taken in the `ValidateReject` action element. The reject handling itself is implemented as a dedicated

`HandleReject` action element. If the retail company agrees with the complaint, it can modify the delivery and state the delivery again. If the executor does not handle the reject, the transaction reaches the end state *Stopped* through the `Stop` action element.

5 DISCUSSION AND CONCLUSIONS

In this paper, it has been shown that an NSBT, which consists only of Normalized Systems elements, can be constructed as a core diagram for an enterprise architecture framework. By using the combination of implementation-independent Enterprise Ontology models and evolvable Normalized Systems elements, we have demonstrated the flexibility in an e-business example, by adding a multi-channel request process. Should the request of the delivery transaction not have been separated from the execution (i.e., the actual delivery), such an addition would have resulted in a combinatorial effect. Consider the following situation, in which this separation has been neglected: separate delivery systems are used for physical requests (i.e., which are requested through mail) and electronical requests (i.e., orders from the website). A simple change in the delivery execution, such as recording the delivery of a certain product, would require two distinct implementations. The impact of the change would thus depend on the size of the system: when more retail channels would be added, the same change would cause an even larger impact. Such combinatorial effects should be avoided. In subsequent iterations, we will integrate other aspects present in enterprise architecture frameworks. This will be done analogously to the integration of cross-cutting concerns on the software level into Normalized Systems elements.

This paper has two important contributions. A first contribution is that we introduced the concept of combinatorial effects on the level of enterprise architectures. We further illustrated how the systems theoretic concept of stability can be applied to the design of enterprise architectures. This requires the elimination of combinatorial effects, which will lead to more evolvable organizations. As a result, we offer a view on enterprise agility that has a strong theoretical foundation. A second contribution is that we demonstrated the feasibility of constructing an enterprise architecture core diagram based on existing scientific approaches. By expressing the core diagram in Normalized Systems elements, we extended the Normalized Systems approach to the organizational level. Using Enterprise Ontology models as the basis for

the core diagram further demonstrates the feasibility of constructing an enterprise architecture framework based on scientific theories. This illustrates how theories from relevant fields can be applied in a new setting by using a design science approach.

REFERENCES

Bernus, P., Nemes, L., and Smidth, G. (2003). *Handbook on Enterprise Architecture*. International Handbooks on Information Systems. Springer-Verlag, Berlin.

Dietz, J. L. (2006). *Enterprise Ontology: Theory and Methodology*. Springer, Berlin.

Hammer, M. (1990). Reengineering work: Don't automate, obliterate. *Harvard Business Review*, 68(4):104ff.

Hoogervorst, J. A. P. (2009). *Enterprise Governance and Enterprise Engineering (The Enterprise Engineering Series)*. Springer, 1st edition.

Lankhorst, M. M. (2005). Enterprise architecture modelling–the issue of integration. *Advanced Engineering Informatics*, 18(4):205 – 216. Enterprise Modelling and System Support.

Lehman, M. (1980). Programs, life cycles, and laws of software evolution. *Proceedings of the IEEE*, 68:1060–1076.

Leist, S. and Zellner, G. (2006). Evaluation of current architecture frameworks. In *SAC '06: Proceedings of the 2006 ACM symposium on Applied computing*, pages 1546–1553, New York, NY, USA. ACM.

Liles, D. H., Johnson, M., Meade, L., and Underdown, D. (1995). Enterprise engineering: A discipline? In for Enterprise Engineering, S., editor, *Proceedings of the Society for Enterprise Engineering Conference*, Kettering, Ohio.

Mannaert, H. and Verelst, J. (2009). *Normalized Systems: Re-creating Information Technology Based on Laws for Software Evolvability*. Koppa, Hasselt.

Ross, J. W., Weill, P., and Robertson., D. C. (2006). *Enterprise Architecture as Strategy – Creating a Foundation for Business Execution*. Harvard Business School Press, Boston, MA.

MODEL-BASED SERVICE INTEGRATION FOR EXTENSIBLE ENTERPRISE SYSTEMS WITH ADAPTATION PATTERNS

Markus Heller and Matthias Allgaier

SAP Research, Vincenz-Priessnitz-Str. 1, 76131 Karlsruhe, Germany
{markus.heller, matthias.allgaier}@sap.com

Abstract: The integration of services into business applications within enterprise services is needed in on-premise settings as well as in upcoming on-demand enterprise settings. Today, it is typically performed in manual integration projects by highly skilled integration experts on the consumer side. As consumers demand flexible and adaptable enterprise systems with lower total cost of ownership (TCO), enterprise system vendors need to provide efficient mechanisms to integrate services within business applications. For this less explored but promising area of service-oriented architecture (SOA) research, a service integration framework with a pattern-based modeling approach is presented that allows for the integration of services into business applications at a later stage in the software-lifecycle - especially after shipment.

1 INTRODUCTION

In the vision of an Internet of Services, organizations in a service ecosystem (Barros and Dumas, 2006) interact in service provider and/or consumer roles to offer, find, trade, and use services like tradable goods. A key challenge for enterprise organisations will be to ensure simple consumption of their offered services via multiple channels (e.g. composite applications, mash-up applications). An important channel handles service consumption within standard business applications running within enterprise systems (e.g. ERP, CRM, or others). Such systems typically realize implemented standard business processes and it is very valuable to be able to extend them with complementary services after the system has been shipped. Such services can be provided by other organizations as service ecosystem partners and offered in service marketplaces or service stores. A less-complex integration of services becomes particular relevant for service/partner ecosystems in upcoming on-demand enterprise software environments (Lo et al., 2009). Two main integration scenarios can be differentiated to achieve the integration of a service into an enterprise system. *First*, existing interfaces of the enterprise system are used to integrate the service. This is possible, if the system provider has foreseen such interfaces at time of shipment. Prominent examples for this first alternative are

standard interfaces for Business-to-Business (B2B) or Application-to-Application (A2A) integration scenarios for Enterprise Application Integration (EAI). *Second*, if a service cannot be integrated into a business application using foreseen interfaces, the enterprise system has to be extended and/or adapted on affected internal application layers beforehand. For example, new UI elements (presentation layer), process steps (process layer), or business object fields (business object layer) can be created or existing elements adapted in the application. Examples are services offered by multiple providers for different core or niche business domains. In both integration scenarios structural- and/or behavioural mismatches between the service interfaces of the enterprise system and the service provider need to be handled by service mediation components.

Three *main deficits* are noted for the second integration scenario: (i) The adaptation or extension of an enterprise system usually requires a high manual effort and high level of expertise. (ii) Today, enterprise systems provide adaptation or extension mechanisms which typically are proprietary solutions on low abstraction level (e.g. code-level interfaces). Similarly, for different application layers typically different adaptation and extension techniques are offered. (iii) Typically, deep domain or technical knowledge is required to oversee and implement the integration solution.

In this paper a service integration framework is proposed for unforeseen service integration into core

business applications with the observed deficits (scenario 2). It enables partners in a service ecosystem to seamlessly integrate new services into business applications within enterprise systems at a later stage in the software-lifecycle. This work applies the design science research methodology (Hevner et al., 2004). Section 2 describes an application scenario from the automotive industry. Section 3 introduces the service integration framework. Related work is described in Section 4. Summary and outlook are given in Section 5.

2 APPLICATION SCENARIO

In the automotive domain, due to legal changes in export guidelines, a manufacturer of car seats needs to certify his products to guarantee that materials used in a car seats are compliant with ecological laws. On a service marketplace, a service provider offers a service for the calculation of eco values for products including certification. The manufacturer runs an enterprise system including a Product-Lifecycle-Management (PLM) module for the design process support for car seats (Figure 1). In its core version, this business application does not support eco value calculation for a given bill of material. Therefore, a product designer of the manufacturer company searches the Service Marketplace directly from within his enterprise system to find services that provide the missing functionality. He receives a list of matching services from various service providers certified for his
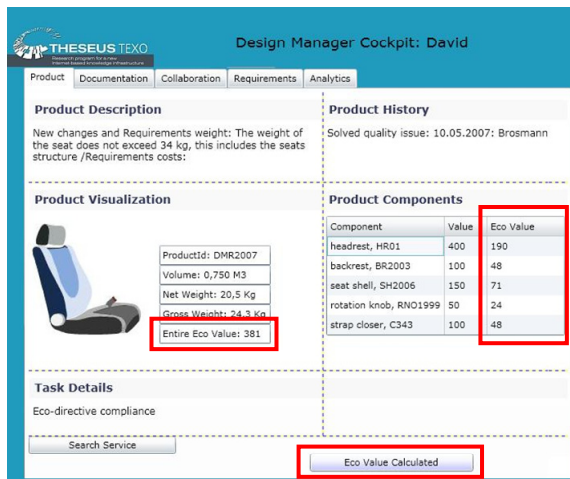
Figure 1: Extension of a PLM business application.

enterprise system. The designer selects and buys a service called "Eco-Calculator" on the marketplace.

This remote service is automatically integrated into the core business application (without the need to run a manual integration project). The user interface of the core business application is extended with (1) a new table column ("Eco Value") in the product components table, (2) a new button ("Calculate Eco Value") and (3) a new field to display the total eco value for the car seat ("Entire Eco Value"). The service is used by the designer immediately after the integration has been completed. If the total eco value fulfils legal requirements, a certificate is generated for the seat and passed to the consumer application.

3 SERVICE INTEGRATION FRAMEWORK

A discussion of the requirements for the framework development is presented in (Allgaier and Heller, 2009). The service integration framework is based on the following characteristics: (1) A *model-based integration approach* is introduced to enable the modeling (or design) of the relevant integration aspects on a higher abstraction level. (2) A *pattern-based modeling approach* is defined that covers typical adaptation- and extension tasks. It allows *controlled extensibility* of the core business applications, insofar, as only a proven set of adaptation operations can be performed. This approach provides a uniform approach to enable service integrators to design the adaptation and extension of a core business application on multiple layers (e.g. UI- and business process layer). (3) To increase the level of automation for service integration, a seamlessly integrated runtime support for extension and/or adaption of enterprise systems completes the framework. This paper describes the (visual) modelling language, while the internal realization of the framework meta model is described in (Allgaier, Heller, and Weidner, 2010).

3.1 Overview and Main Components

Figure 2 shows the main components of the service integration framework: Integration Modeling Environment (top), Adaptation-/Extension Exe-cution Environment (middle), and Enterprise System (bottom). (1) As a prerequisite, the enter-prise system provider creates and delivers an Enterprise System to run different Core Business Applications. These business applications are used by end users. The enterprise system is connected to the Adaptation/Enactment Execution Environment to allow for adaptation and extension of the system after initial shipment at a later stage in the software-

lifecycle. The enterprise system provider publishes an Application Extensibility Description (left) that represents the core business application's extensibility capabilities within the enterprise system. Similarly, the service provider creates a Service Description (right) for a service that he offers, e.g. on a service marketplace. The Service Description describes multiple aspects of the service's capabilities.

(2) The service integrator uses the Integration Modeling Environment to model all relevant aspects to define an integration of a selected service into a selected core business application within an enterprise system (based on a loaded application extensibility description and a loaded service description). All modeled information about the integration is stored in an Integration Model data structure within the modeling environment. This model specifies all adaptation and/or extension steps to achieve the desired integration.

(3) When all details have been completely specified, the Integration Modeling Environment creates an Integration Description (middle) that can be loaded into the Adaptation/Extension Execution Environment. This component orchestrates all necessary steps for the adaptation and extension of the core business applications and enterprise system.

An Enterprise System (e.g. ERP system) is described with a single overall abstracted model that spans across four abstraction layers. Basically an enterprise system consists of multiple (service-based) Business Applications that leverage a common service and business configuration layer. The Service Layer contains all services offered by the enterprise system. Core services provide access to business objects and they can be composed into larger bundles to provide advanced higher-value business functionality or application logic. The Business Configuration Layer contains the configuration data for the business applications with all of its available customization options.

For a business application, the Presentation Layer comprises all artifacts and components for the user interface part of the business application (for example, UI components for a dedicated UI platform with all interrelations). Likewise, the Process Layer contains models of the business processes which are realized within a business application. Modeling elements for business processes can contain references to elements on other layers such as presentation or service layer.

To adapt standard business applications to customer specific needs, enterprise systems typically provide a large set of proprietary extensibility resp. adaptability features (as e.g. in SAP's ByDesign On-Demand platform). The features support a wide
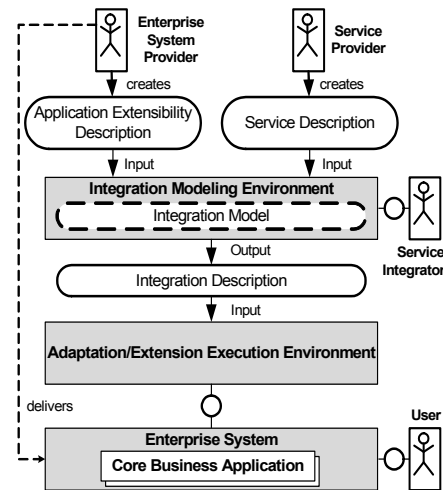
Figure 2: Service Integration Framework Components.

spectrum of use cases and address various stakeholders for flexibility requirements like customers (e.g. extensibility or flexibility as customer self-service), verticalization and globalization, or partners in software eco systems.

3.2 Layered Modeling Approach

The mentioned central model artefacts in the framework are detailed as follows: (a) The *Application Description* contains all possible extension points where the application can be extended or adapted. Such extension points model the offered extensibility features of the application on top of the underlying enterprise system. For example, extension points denote places in the application that can be used to add a process step to a core process or a new UI, and so on. (b) The *Service Description* contains meta-data information about the service and all information needed for its integration into business applications (such as service operations with input and output interfaces, supported data types, messaging choreographies, offered default UI descriptions). (c) The *Integration Description* contains as a dominant part a set of connections between elements the application extensibility description and the service description as well as additional parameterization data. It can additionally reference any combinations of software artifacts needed for the Adaptation/Extension Execution Environment to perform the service integration.

Application description, integration description, and service description (shown horizontally in Figure 3 in the columns from left to right) are modeled on two different modeling layers (shown vertically). The *Technical Model Layer* (bottom

165

part) contains models to describe applications or services with respect to their extensibility support. Standard Models (like e.g. XAML, BPMN) lack necessary modeling elements (for example, to model extension points). Therefore, they need to be enhanced on the technical model layer with new extensibility concepts. The Enhanced Application Model represents a complete description of an extensible application. It comprises all kinds of models describing the extensibility support of the application on different application layers. Presentation or process models (XAML, BPM, etc.) are *enhanced* with additional new elements to model extension points and extension connection models (e.g. as in the BPMN 2.0 proposal draft). The Enriched Service Model (e.g. in USDL) contains modeling elements from a service model (like WSDL) plus new elements to describe further service capabilities like supported business application details. The Technical Integration Model holds all connections between elements from Enhanced Application Model and Enriched Service Model with parameterizations.

The *Semantic Model Layer* (top part) contains ontology representations of the models from the Technical Integration Layer. For this purpose, relevant elements from models on the Technical Model Layer are transformed into ontological representations on the Semantic Model Layer. A Semantic Application Model semantically captures important model concepts from the Technical Application Model (with extensibility support). Likewise, the Semantic Service Model and the Semantic Integration Model are constructed from the models on the Technical Layer.
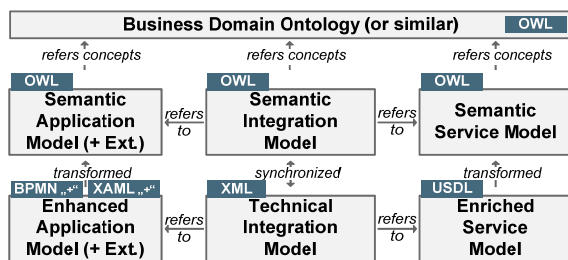
Figure 3: Technical and semantic modeling layers.

Some of the models on both layers can reference entities of further ontologies (top part) that model important business domain concepts, like in Business Domain Ontologies (BDO), e.g. from the SUPER research project. Within the framework, the Technical Models (as Ecore models) are used together with the Semantic Models (as OWL ontologies) which primarily serve for reasoning over the model semantics for searches in model data and advanced modeling guidance support. In the rest of this paper, the semantic models are used. They are valuable for advanced semantic search functionality to recommend best fitting application extension points or service elements for ports of used pattern instances in the modeling environment.

3.3 Pattern-based Modeling

The integration modeling language is based on the notion of adaptation pattern. They allow to restrict which integration steps are performed to control the flexibility of the enterprise system. (Atomic) Adaptation Patterns define typical fine-grained adaptation or extension tasks that can be performed on the application layers. Adaptation Patterns represent parameterizable connection links between extension points of the application description and elements of the service description. The patterns have a set of application or service reference ports and a set of further parameterization attributes. An application reference port links the pattern to an application extension point (of a predefined type), while a service reference port links the pattern to a service description element (of a predefined type). Parameterization attributes store further key-value pairs within the pattern context.

Complex adaptation patterns are composed of atomic and/or other complex adaptation patterns and they model more coarse-grained extension tasks for typical multi-step integration scenarios, e.g. across different application layers. Atomic and composite patterns are stored in a knowledge base called *adaptation pattern catalog* and they can be reused within many integration designs by different service integrators (best-practice sharing). The set of stored adaptation patterns can be collected, extended, and revised based on other related work, e.g. (Weber, 2008) for possible pattern candidates or by creating previously unknown patterns (e.g. an adaptation pattern editor tool is currently under development). During the modeling phase, an integration model instance is created in the modeling environment by the service integrator and extended step-wise by first choosing an (complex or atomic) adaptation pattern from a catalog and adding it to the current integration model. Second, the pattern's ports are connected to chosen application extension points or service description elements.

An *example* from the Eco Calculator scenario in Section 2 illustrates the modeling approach.

The service integrator wants to define the integration of the Eco Calculator service (found on the service marketplace) into a business application.
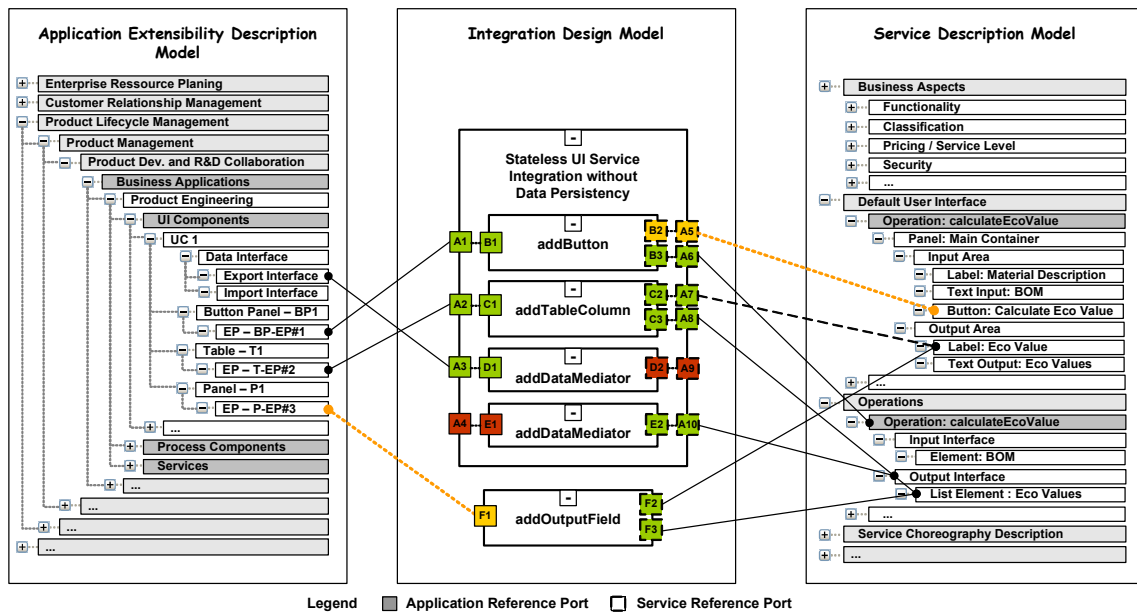
Figure 4: Example Integration Model for the Integration of the EcoCalculator service into the PLM business application.

No interface was foreseen for this use case in the enterprise system at shipping time. The service integrator identifies the following requirements (e.g. from a business department): The service should be integrated into the Product-Lifecycle-Management (PLM) part of the enterprise system. The service should be used before the product (car seat) is shipped. The eco values returned from the service should be displayed on the user interface in the existing table to display the bill of material elements for the car seat. The returned eco values should only be displayed on the user interface.

Figure 4 shows integration model (middle), application description (left) for the PLM application and service description (right) for the EcoCalculator service. The model contains one complex adaptation pattern instance and another atomic adaptation pattern instance with their parameterizations. The complex pattern "Stateless Service Integration without data persistency" was chosen because it technically seems to fit to the given business requirements to integrate the service on UI layer and service layer: It allows for adding a stateless service (with one operation), adding a UI button and display labels into UI panels. The result data of the service is not persisted in the enterprise system. The modeled complex pattern contains four atomic adaptation patterns: (i) addButton, (ii) addTableColumn, (iii) addDataMeditator and (iv) addDataMediator. The complex adaptation pattern carries four application reference ports (A1, A2, A3, and A4) and six service reference ports (A5, A6, A7, A8, A9, and A10) that are internally linked to the

ports of the contained atomic adaptation patterns. Some ports are parameterized with application's extension points or service description elements. For example, for the chosen application "Product Engineering", some extension points of the UI component "UC1" are connected to the application ports (A1, A2, A3, and A4) of the shown adaptation patterns and some service elements are connected to the service ports (A5, A6, A7, A8, A9, and A10).

The adaptation pattern "addButton" is connected via the application reference port B1 with the port type *Extension Point Type – Button Panel*. This port is parameterized with the value *BP-EP#1*. The text for the button is taken from the *Default User Interface* section of the service description (service reference port B2). The information for the button's event handler (service operation that is called when the button is pressed) is taken from the *Operations* section of the service description (service reference port B3). The pattern "addDataMediator" models a data mediation problem which is resolved via externally executed data mediation tools. Finally, a new integration description (referencing application description and service description) is generated from the modeling environment.

3.4 Runtime Extensibility Support

An implemented prototype demonstrates the feasibility of the proposed approach to service integration for the application scenario of Section 3. The PLM business application is implemented based on Microsoft Silverlight, the EcoCalculator Web

Service uses the AXIS framework. The prototype addresses the presentation layer. Figure 5 shows its architecture. An integration description is loaded into the Adaptation/Extension Execution Environment (Java). It analyzes each adaptation step in the description and forwards it to a layer-specific adaptation manager, here for UI- and service-layer. They adapt the PLM application, for example, for the UI-layer adaptation by calling a sequence of commands at the native extensibility features API (MS Silverlight API) at runtime. The screenshot in Figure 1 on page 4 shows the prototype after extension with a new button, a new table column, and a new output label. Not visible in the UI, two data mediators are added to map data between the UI context and the service interface.

4 RELATED WORK

This work addresses the controlled extensibility of enterprise systems for unforeseen service integration, similar to related B2B Integration and Enterprise Application Integration, e.g. (Hophe and Woolf, 2003). Structural or behavioural interface mediation techniques, e.g. (Studer et al., 2007), are leveraged in the framework, e.g. for data mediation. Work on adaptive software systems typically addresses self-adaptive systems for mobile, pervasive computing (e.g. MADAM, http://www.ist-madam.org). Plug-in techniques for development and installation of (downloaded) components into component-based core application frameworks, e.g. (Birsan, 2005), or for runtime adaptation of ERP systems, e.g. (Wolfinger et al., 2008), do not target unforeseen service integration. Extensions on the presentation layer leverage adaptive user interface modeling approaches (e.g. XIML, UIML). Process change patterns, e.g. by (Weber et al., 2008), provide a conceptual basis for the process layer adaptation pattern catalog in this paper.

5 SUMMARY AND OUTLOOK

A model-based service integration framework for the unforeseen integration of services into extensible enterprise systems has been presented. In the author's opinion, the promising service integration area should be further investigated. The modeling approach with adaptation patterns and runtime support is demonstrated with a UI integration prototype in the automotive domain. The Process layer support is currently developed.
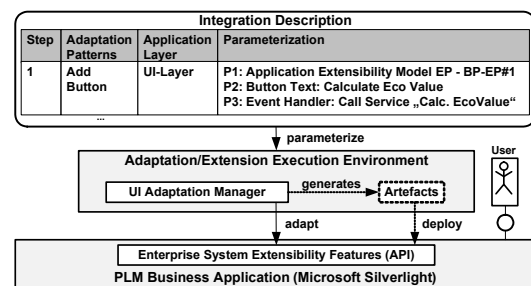
Figure 5: Runtime support architecture.

ACKNOWLEDGEMENTS

REFERENCES

Allgaier M. and Heller M. (2009) *Research Challenges for Seamless Service Integration in Extensible Enterprise Systems*, Workshop "Industrial Experiences for Service Oriented Computing". Stockholm, Sweden.

Allgaier M., Heller M., and Weidner M. (2010) *Towards a Model-based Service Integration Framework for Extensible Enterprise Systems*, in: M. Schumann, L. M. Kolbe, M. H. Breitner (Eds.): Tagungsband der Multikonferenz Wirtschaftsinformatik, Göttingen.

Barros A. and Dumas M. (2006) *The Rise of Web Service Ecosystems*. IT Professional, 8, 5, 31-37.

Birsan D. (2005) *On Plug-ins and Extensible Architetures*. Queue, 3, 2, 40-46.

Hevner A.R., March S.T., Park J., and Ram S. (2004) *Design Science in Information Systems Research*. MIS Quarterly, 28, 1, 75-105.

Hohpe G. and Woolf B. (2003) *Enterprise Integration Patterns - Designing, Building, and Deploying Messaging Solutions*. Addison-Wesley Prof., Boston.

Lo H., Wang R., and Garbini, J.P. (2009) *The State of Enterprise Software 2009*. Forrester, Cambridge.

Studer R., Grimm S., and Abecker A. (2007) Semantic Web Services. Springer, Berlin.

Weber B., Reichert M., and Rinderle-Ma S. (2008) *Change patterns and change support features - Enhancing flexibility*. DKE, 66, 3, 438-466.

Wolfinger R., Reiter S., Dhungana D., Grünbacher P., and Prähofer H. (2008) *Supporting Runtime System Adaptation through Product Line Engineering and Plug-in Techniques*. In: 7th IEEE International Conference on Composition-Based Software Systems (ICCBSS'08), Madrid.

TOWARDS ONTOLOGY BASED E-BUSINESS STANDARDS

Bahareh Rahmanzadeh Heravi, David Bell, Mark Lycett

Department of Information Systems and Computing, Brunel University, Uxbridge, London, UB8 3PH, U.K.
{Bahareh.Heravi, David.Bell, Mark.Lycett}@brunel.ac.uk

Stephen D. Green

Document Engineering Services, Cambridgeshire, CB6 2AP, U.K.
stephen.green@documentengineeringservices.com

David Snelling

Fujitsu Laboratories of Europe Ltd., Hayes Park Central, Hayes, Middlesex, UB4 8FE, U.K.
david.snelling@uk.fujitsu.com

Abstract: e-Business standards are recognised as one of the most important drivers of Business to Business Integration. These standards seek to provide unambiguous specifications for error-free exchange of documents and information between trading partners. These standards are however, syntax based and do not guarantee semantic interoperability between partners. This paper proposes the utilisation of semantic web technologies in the standards development process, aiming at developing more robust and at the same time flexible e-Business standard. In order to extract the requirements of an ontology based standard, a combined top down and bottom up approach has been adopted. This resulted in developing two ontologies: one for e-Business standards in general and one for ebXML Business Process Specification Schema (ebBP) as a specific e-Business standard. The challenge is to address the distance between these two ontologies and explore how ontologies can be utilised in developing next generation e-Business standards. It is believed that ontology based e-Business standards will enhance interoperability between organisations involved in value networks and also may facilitate the standard development process itself.

1 INTRODUCTION

Inter-organisational collaborations are effective means of gaining competitive advantage and improving effectiveness for today's organisations. In such an environment, collaborating parties' business processes and their associated documents need to be understood and aligned across organisational boundaries. E-business standards are traditionally used for achieving interoperability between trading partners and aimed at error-free exchange of documents and information. E-business standards, however, only provide syntactic and not semantic interoperation, since they are mainly based on XML, whose provision for semantic knowledge sharing is particularly restricted.

Singh, Iyer and Salam (2005a) provide a vision for Semantic e-Business which is based on Tim Berners-Lee's Semantic Web vision (Berners-Lee, Hendler & Lassila, 2001). In this vision, semantic e-Business is introduced as an approach to managing knowledge for the coordination of e-Business processes through the systematic application of Semantic Web technologies (Singh, Iyer & Salam, 2005b). They argue that Semantic e-Business will be enhanced through more rigorous information and knowledge exchange. Ontologies can capture the definitions and interrelationships of concepts in a variety of domains, resulting in a shared understanding of the domain, which is indeed the ultimate goal of e-Business standards.

This paper explores the intersection of the e-Business standards and the ontologies, with the aim of improving e-Business standards, which in turn better facilitate B2B integration. This can have a significant effect on research community as well as practitioners, who are involved in value networks.

The remainder of this paper is structured as follows: Section 2 provides a background on e-Business standards and the structure of standards in general. Section 3 explores the use of ontologies in eBusiness standards. Section 4 introduces the concept of ontology based e-Business standards and its benefits, provides a methodology for deriving an ontology for e-Business standards and presents a work in progress ontology on that basis. Section 5 presents the challenges associated with utilising ontologies as a basis for e-Business standards, followed by a discussion on the next steps of the research and a set of research questions. Section 6 concludes the paper.

2 E-BUSINESS STANDARDS

A standard is a technical specification approved by a recognised standardisation body, which is designed to be used consistently, as a rule, a guideline, or a definition across particular communities of interest, to achieve mutual benefit (ETSI, 2010). Acquiring 'standard' status may take several years. During this time a specification may be implemented if it receives sufficient public review and achieves a certain level of approval. The specification's use may be widespread even without full standardisation. Openness of standards / specifications is usually considered as an important factor in achieving consensus and widespread adoption. When agreed as a 'standard', it may only be the status of the specification which changes.

Each specification / standard is composed of a set of Normative Statements, often with a Conformance Clause and associated Test Assertions.

A Normative Statement defines the prescriptive requirements on a conformance target (Green, Kostovarov, 2009). In the standardisation terminology conformance refers to the fulfilment of specified requirements by an implementation of the standard. This is verified with the use of Conformance Clauses, which must, directly or indirectly, reference one or more Normative Statements and may also refer to another Conformance Clause (Green, Kostovarov, 2009).

Furthermore, a Test Assertion is an independent, complete, testable or measurable statement for evaluating the adherence of part of an implementation to a Normative Statement in a specification (OASIS TAG TC, 2010)(Durand, Green, Kulvatunyou, Rutt, 2009).

E-Business standards seek to establish Interoperability between trading partners by defining standard interfaces specifying one or more common Business Processes, elements of the Business Documents and / or Messaging details. Figure 1 presents an Interoperability Stack for e-Business standards and specifications. E-business specifications may cover one or more of the layers in the stack.
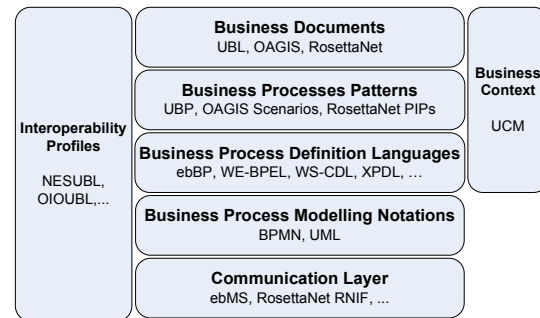
Figure 1: e-Business Interoperability Stack.

The Communication layer provides specifications for packaging, security and transport of messages to be used within business interactions. ebXML Messaging Service (ebMS) and RosettaNet Implementation Framework (RNIF) are examples of specifications from this layer.

Business Process Modelling Notations are specifications which are used for graphical representation of processes and do not have XML representation. Business Process Modelling Notation (BPMN) and UML Activity diagrams are examples from this layer.

Business Process Definition Languages provide specifications for XML based representation of business processes. These languages can have different targets. For example they might be suitable for public choreographies, such as ebXML Business Process Specification Schema (ebBP), for business process execution, such as WS-BPEL or for private workflows such as XPDL. These processes may be visualised by Business Process Modelling Notations.

The next level, Business Process Patterns, provide specifications for repeatable processes which need to be agreed between trading partners, such as procurement processes. These processes can be industry neutral or industry specific depending on the target of the specification. UBP (Universal Business Process) and OAGIS (Open Applications Group Integration Specification) Scenarios are examples of industry neutral processes and RosettaNet PIPs are examples of industry specific business processes. These processes may be represented by Business Process Definition

Languages and / or visualised by Business Process Modelling Notations.

B2B transactions are composed of Business Document exchanges, within the steps of Business Processes. Therefore, the next layer represents Business Document specifications, such as UBL (Universal Business Language) and xCBL (XML Common Business Library. Business Document standards can also be industry specific or industry neutral.

Business Context is another layer in the e-Business Interoperability stack and provides contextual information to be used in Business Documents and Business Process specifications. Unified Context Methodology (UCM) is an example from this layer, which is a UN/CEFACT specification and aims at facilitating context-sensitive modelling of e-Business transactions.

Interoperability Profiles are subsets of standard specifications which focus on specific business processes or industries. Northern European Section UBL (NESUBL) is an example from this layer.

Currently e-Business standards are mainly based on XML. Built upon W3C standards, XML based e-Business standards, such as ebXML and RosettaNet, provide a good basis for a common syntactical understanding between trading partners. XML based e-Business standards are a big step towards B2B integration and have been quite successful in providing general and well utilised syntactic standards. However, they cannot facilitate semantic integration between business partners as XML can only cover syntax and not the semantics of the transactions.

Ontologies, on the other hand, are an appropriate means of unambiguously capturing the definitions and interrelationships of concepts in a formal, unambiguous and machine interpretable manner, with the aim of a shared understanding of a domain, which is indeed the ultimate goal of e-Business standards. Therefore, utilising ontologies seems an appropriate approach for defining more expressive, stable and interoperable e-Business standards.

3 E-BUSINESS STANDARDS AND ONTOLOGIES

A considerable number of publications emphasise on the importance of semantic web technologies and ontologies in B2B transactions (Legner, Wende, 2007)(Kajan, Stoimenov, 2005)(Wu, Li & Yang, 2006)(Gong, Ning, Chen, O'Sullivan,

2006)(Höfferer, 2007)(Liegl, Huemer & Zapletal 2009)(Vujasinovic et al., 2010). There are also a growing number of ontologies developed for e-Business related standards in the literature. Examples are oXPDL, an ontology for XPDL (Haller, Gaaloul & Marmolowski, 2008), an ontology for WS-BPEL (Nitzsche, Wutke & Van Lessen, 2007), ebXML Registry Profile for OWL (OASIS ebXML Registry TC, 2006), which provides specifications for publishing and discovering OWL ontologies in the ebXML Registry/Repository and OntologUBL, which provides an ontology for Universal Business Language (The Ontolog Forum, 2002).

There are also a few works focusing on utilising ontologies in conjunction with e-Business standards. Vujasinovic, Ivezic, Kulvatunyou, Barkmeyer, Missikof, Marjanovic and Miletic (2010) provide a semantic mediation architecture for standard based B2B interoperability. This work emphasises the importance of Standard Development Organisations in achieving standard based semantic B2B integration and thus highlights the importance of ontologies in relation with e-Business standards. OASIS may be considered as the first Standard Development Organisation to address ontologies and semantic web technologies and their synergy with standards. The first ontology related initiative in OASIS is the Semantic Support for Electronic Business Document Interoperability Technical Committee (OASIS SET TC, 2009), which aims at developing specifications for machine processable semantic content of the Electronic Business Documents based on the UN/CEFACT Core Components Technical Specification (CCTS). Another relevant TC in OASIS, which may be considered as the first official ontology based standard Technical Committee, is called OASIS Quantities and Units of Measure Ontology Standard (QUOMOS) Technical Committee (OASIS QUOMOS TC, 2010). Ontolog forum (The Ontolog Community, 2010) is another relevant initiative which addresses the importance of ontologies for standard community and therefore had 'Toward Ontology-based Standards' as their 2009 ontology summit theme. In fact OASIS QUOMOS was the result of discussions in the ontolog forum, which ended up as an OASIS TC. These efforts emphasise on the significance of ontologies and semantic web technologies in the standards world and imply that it is time for the intersection of these two communities.

Nevertheless, almost no effort has yet been taken to utilising ontologies for developing, authoring or

improving e-Business standards. It is believed that this is an important gap, which needs to be explored extensively.

4 ONTOLOGY BASED E-BUSINESS STANDARDS

Ontologies have the potential to facilitate both the creation and utilisation of standards (The Ontolog Community, 2010). They may also be used to improve the quality of standards, leading to more robust implementations as well as the semantic integration of different standards.

Ontologies may contribute to the development, extension and improvement of e-Business standards specifications in the following ways:

1. *Formalise concepts within existing e-Business standards*, such as ebXML and RosettaNet. This would result in a more stable definition of semantics in the standard and allow the writing expressions based on clear, unambiguous terms and categories. In this approach, specifications serve as a foundation for developed ontologies and therefore the ontologies can be evaluated against the standards specifications and their XML Schemas. (The Ontolog Community, 2010)

2. *Reengineering of existing standards based on ontological analysis*, identifying their potential problems and semantic ambiguities and improving them. (The Ontolog Community, 2010)

3. *Facilitate integration between different standards*, which are already defined using ontologies.

4. *Development of standards*, wherein ontologies are used throughout the standard development phases, from start to finish, realising the benefits of the semantic vision outlined in Section 1. This approach can be taken where appropriate in developing new e-Business standards or new versions of existing standards.

The latest of the above is the proposition of this paper and will be further discussed in the remainder of the paper.

In order to study the domain under discussion, two approaches have been adopted: top down and bottom up. The top down approach is done in collaboration with standards developers and experts in the field of standardisation and also taking into account various relevant specifications. The purpose of this approach is to formalise the structure of e-Business standards in a robust and unambiguous way. The output of this approach is an ontology for e-Business standards, which represents the building blocks of standards and their relationships in general and e-Business standards in particular. This ontology may further be utilised as a basis for developing ontology based e-Business standards and possibly other information systems related standards.

The Bottom up approach is taken as an experiment for developing an ontology for a specific e-Business standard, OASIS ebXML Business Process Specification Schema (ebBP). The purpose of the bottom up approach is to explore the requirement of an e-Business standard, in particular, and challenges associated with that. The result of the bottom up approach can also be used to reflect on the top down approach and its resulting ontology.

While conducting the top down approach, a methodology comprising nine steps was developed. Steps 1 to 4 have already been taken and steps 5 to 9 are to be done in the future. The steps are as follows:

Step 1. Brainstorming for developing a domain model.

Step 2. Turning the model into a concrete ontology.

Step 3. Identifying those parts of the model related to the artefacts which are wished to be standardise.

Step 4. Adding related properties to flesh out those parts of the ontology identified in step 3.

Figure 2 depicts a snapshot of the ontology resulting from the methodology up to step 4. It is important to note that this ontology is a work in progress and is solely provided as an introduction of the domain under study. Furthermore, figure 2 depicts a part of the ontology and doesn't include all classes and properties of the ontology.

Steps 5 to 9 will be addressed in future to enhance the ontology are as follows:

Step 5. Add rules to turn the parts of the ontology identified in step 3 into artefacts such as mark-up wished to be standardise.

Step 6. Express the above rules as subject and predicate with identifiers.

Step 7. Use the aforementioned rules to generate the artefacts needed, such as a schema for a mark-up representation of the ontological model.

Step 8. Add test assertions based on the model in step 4 and statements in step 6.

Step 9. Publish the ontology including the model and its representational artefacts, possibly in separate specifications for modularity.

For the bottom up approach an ontology for ebBP is developed. Figure 3 presents a snapshot of a set of classes in the ebBP ontology.
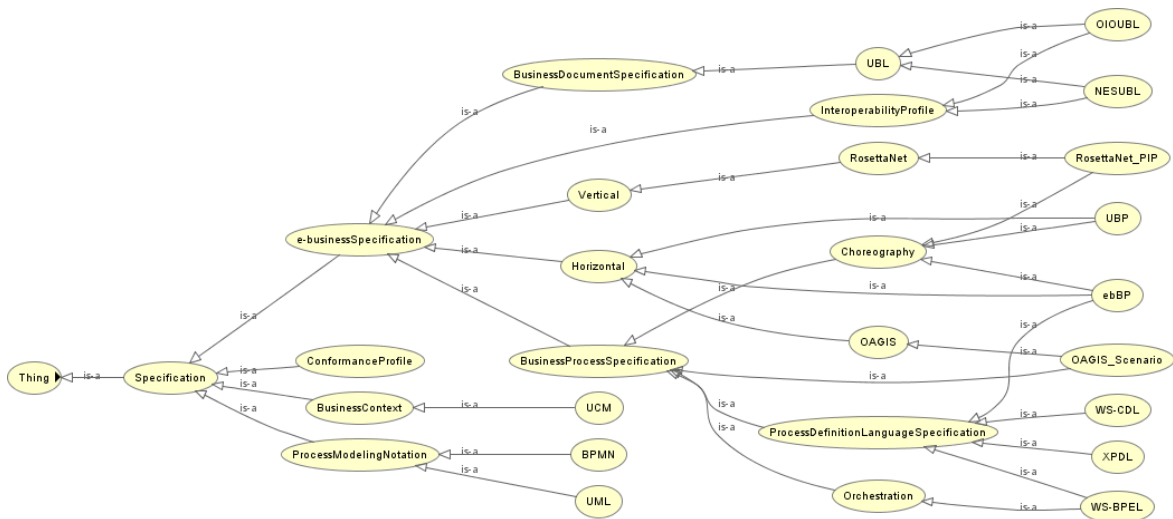
Figure 2: e-Business Specifications Ontology.
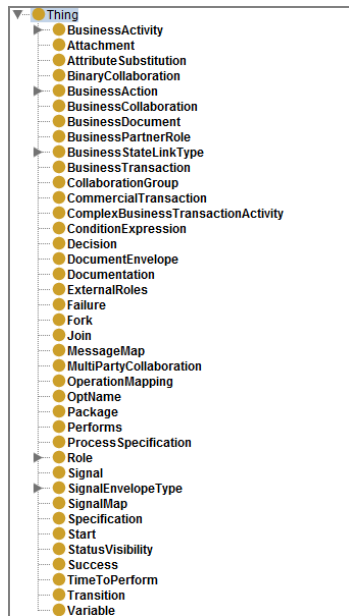
Figure 3: ebXML Business Process Specification Schema Ontology.

In combining the two approaches it seems that there is not a clear correspondence between the classes in the e-Business Specifications ontology and the ebBP ontology. These ambiguities imply that future work is needed in harmonising these two levels on standard ontologies.

This paper suggests that ontologies can be used for defining new e-Business specifications or new version of existing standards in various extents, including their conformance clause, test assertions and normative statements. They can also be used for

conformance checking of the implementations of standards. To achieve this however, a consistent standard development methodology, which utilises ontologies in the whole standard development process, is required.

Using these ontologies in this way, a basis for specifying standards is provided. All or part of a standard may thus be developed from beginning to end making full use of ontologies and also domain-specific ontologies to achieve the benefits of reduced ambiguity and vagueness and to allow the development of complex and dynamic standards with context-dependant rules.

5 CHALLENGES AND FUTURE RESEARCH

5.1 Challenges

There remains the need for the standards technical committee or working group to consider how the use of ontologies as a basis for their specifications affects the implementers and, in some cases, even the end users of their standard. It may be that XML Schema files or other artefacts are to be generated using specialised tools from the ontologies and that such a schema is to define an XML instance for use in a particular context, as specified in context-specific information in the ontologies. It might be necessary to deduce certain requirements by the introduction of values for variables included in the ontologies, normative statements or test assertions; values not known until this stage but accounted for

in the standard. These techniques allow for greater flexibility and wider scope in defining a standard for implementation in a variety of situations. In some cases it may be expected that implementations or end user software will include knowledge bases and other ontology-aware technologies to make full use of the ontology basis of the standard. Test cases may be stored and retrieved dynamically from a database or repository according to results obtained running previous tests and according to complex rules based on ontology-based test assertions.

Challenges abound as the benefits do not offer themselves without a struggle. The obvious but often neglected challenge is to provide change management and version control facilities both for the standard and for the implementations of the standard.

Nevertheless, despite the complexities and stringent requirements on both standard developer and standard implementer, ontologies are becoming more and more a common sight in standards committees as their benefits are recognised for their potential to improve quality or implementation and interoperability.

5.2 Next Steps and Research Questions

As mentioned in section 4, the top down ontology provided in this paper is a works in progress, which needs to be extended and fully tested in different cases and in relation with different standards. In the next steps of these research steps 5 to 9 explained in section 4 will be taken to complete the e-Business standards ontology. The plan is to test the resulting ontology with a number of ontologies for ebXML standards, which will be defined using the bottom up approach, and explore the outcome of the test.

When considering utilising ontologies as a basis for e-Business standard development a number of research question require consideration:

- How normative statements can be defined using an ontology based e-Business standard?
- How test assertions can be addressed in an ontology based e-Business standard?
- How conformance clauses can be defined and represented in the ontology based e-Business standards and how they can be queried for conformance checking of a specific implementation?
- How existing standards can be included in the ontology based framework?
- How different versions of a standard can be managed using this approach?

- How to relate a model to its representations and how to relate the representation ontology classes to external representation definitions such as those in a schema?

6 CONCLUSIONS

This paper proposes a novel approach in developing e-Business standards and suggests that ontologies should be used in the process of e-Business standards development in order to fully realise the semantic e-Business vision. To explore this idea a current state of the art is provided followed by a combined top down and bottom up approach adopted to develop an ontology for e-Business standards in general and an ontology for a specific e-Business standard, respectively. Analysing the state of the art and comparison between top down and bottom up approaches suggest that more research is required in this field and therefore the paper is concluded with a research agenda and relevant research questions to be addressed in the future. To conclude, it is argued that development of standards, which are based on ontologies, would enhance their stability and usability, as well as facilitating standard based integration and interoperation in value chains.

ACKNOWLEDGEMENTS

REFERENCES

Berners-Lee, T., Hendler, J. & Lassila, O. 2001, "The Semantic Web", vol. 284(5), pp. 34-43.

Durand, J., Green, S.D., Kulvatunyou, S. & Rutt, T. "Test Assertions on steroids for XML artifacts", *In Proceedings of Balisage: The Markup Conference 2009. Balisage Series on Markup Technologies, vol. 3 (2009).*

European Telecommunications Standards Institute (ETSI). Retrieved February 26, 2010, from http://www.etsi.org

Gong, R., Li, Q., Ning, K., Chen, Y. & O'Sullivan, D. 2006, *Business process collaboration using semantic interoperability: Review and framework.*

Green, S.D. & Kostovarov, D. "Test Assertions Guidelines Version 1.0", 2009.

Haller, A., Gaaloul, W. & Marmolowski, M. 2008, "Towards an XPDL compliant process ontology", pp. 83.

Höfferer, P. 2007, "Achieving Business Process Model Interoperability Using Metamodels amd Ontologies", *Proc.of the 15th European Conference on Information Systems (ECIS2007),* , pp. 1620-1631.

Kajan, E. & Stoimenov, L. 2005, "Toward an ontology-driven architectural framework for B2B", *Communications of the ACM,* vol. 48, no. 12, pp. 60-66.

Legner, C. & Wende, K. 2007, "The Challenges of inter-organizational Business Process Design - A research Agenda", *15th European Conference on Information Systems* , pp. 106-118.

Liegl, P., Huemer, C. & Zapletal, M. 2009, "Towards a global business document reference ontology", pp. 355.

Nitzsche, J., Wutke, D. & Van Lessen, T. 2007, "An ontology for executable business processes", *Workshop SBPM,* vol. 251, pp. 52-63.

OASIS ebXML Registry TC, 2006, *ebXML Registry Profile for Web Ontology Language (OWL) Version 1.5*. Retrieved February 26, 2010, from http://xml.coverpages.org/OASIS-regrepOWL-Profile-22611.pdf

OASIS QUOMOS TC (2010). Retrieved February 26, 2010, from http://www.oasis-open.org/committees/ tc_home.php?wg_abbrev=quomos.

OASIS SET TC. (2009). Retrieved February 26, 2010, from http://www.oasis-open.org/committees/ tc_home.php?wg_abbrev=set.

OASIS TAG TC. (2010). Retrieved February 26, 2010, from http://www.oasis-open.org/committees/ tc_home.php?wg_abbrev=tag

Singh R., Iyer L.R., Salam A.F. 2005a, "The semantic E-business vision", *Communications of the ACM,* vol. 48, no. 12, pp. 38-41.

Singh, R., Iyer, L.S. & Salam, A.F. 2005b, "Semantic eBusiness", *International Journal on Semantic Web & Information Systems,* vol. 1, no. 1, pp. 19-35.

The British Standards Institution. (2010). Retrieved February 26, 2010, from http://www.bsigroup.com

The Ontolog Community. (2010). Retrieved February 26, 2010, from http://ontolog.cim3.net/.

The Ontolog Forum *, UBL Ontolog*y, (2002). Retrieved February 26, 2010, from http://ontolog.cim3.net/cgi-bin/wiki.pl?UblOntology

Vujasinovic, M., Ivezic, N., Kulvatunyou, B., Barkmeyer, E., Missikoff, M., Taglino, F., Marjanovic, Z. & Miletic, I. 2010, "Semantic mediation for standard-based B2B interoperability", *IEEE Internet Computing,* vol. 14, no. 1, pp. 52-63.

Wu, B., Li, L. & Yang, Y. 2006, "Ontological approach towards E-business process automation", pp. 154.

POSTERS

BUSINESS PROCESS FLEXIBILITY
Evaluation Criteria and Guidelines

Christina Tsagkani

University of Athens, Department of Informatics & Telecommunications, Panepistimiopolis, Ilisia, Athens, Greece
tsagkani@di.uoa.gr

Abstract: The business environment of most enterprises comprises of fluid requirements, and emergent behaviour that cause continuous changes across the enterprises' business processes. Thus, Process Management Systems (PMSs) able to handle such changes become a necessity for businesses in order to effectively respond in this volatile environment. However, despite the plethora of available PMS, dynamic process change is hardly being addressed in most of them. Therefore, the task of selecting a PMS that supports flexible business processes effectively and in this way face the volatile nature of the business environment is not easy. This task is being addressed in this paper by proposing a set of evaluation criteria for flexible PMS. In addition, a business case scenario from the banking sector and selection guidelines have been employed, in order to demonstrate how the proposed criteria framework may be applied practically during the selection of the 'best-fit' PMS.

1 INTRODUCTION

The process orientation of contemporary information systems has led to the development of a plethora of Process Management Systems (PMS) making it increasingly difficult for an organization to choose the PMS that is best suited for its own needs. Consequently, quite often we stumble upon organizations which have ended up using more than one PMS suites, limiting the scope of each one to specific business functions or departments. This fact may be the source of various issues like incompatibility between different systems, lack of know-how, limited reusability of specific business process steps, high maintenance costs, etc. that overall hinder the benefits of using an advanced PMS.

Besides, the design phase in a business process lifecycle is addressed by traditional Process Management Systems (PMSs) in a way that provides a static business process incorporating all possible exceptional situations and process extensions. This is hard to achieve, time consuming, and may lead to complex processes. As a result, many PMSs end up being insufficient in today's volatile business environment. Thus, modern PMSs need to include appropriate techniques that support deviations from the original process definition.

Based on the above realities we come to the conclusion that selecting a PMS that supports flexible business processes effectively is not an easy task. This paper tries to contribute in this situation in two ways:

a) Firstly by proposing a set of functional and non-functional evaluation criteria for techniques designed to enact, manage and support flexible business processes. These criteria have been derived based on our experience while working with currently available PMSs that support some kind of process flexibility, analysis of the different features supported by them, as well as extensive literature review.

b) Secondly, by demonstrating how the evaluation criteria could be used during flexible PMS selection. For this purpose we follow some simplified guidelines and use a specific case coming from the banking sector which is our area of expertise. Our aim is to show in practice how a specific organization may be facilitated using the proposed evaluation criteria while selecting a flexible PMS.

The presentation of the results of this work is organized as follows: a brief introduction to flexible business processes is provided in the following section. Next, in Section 3 a set of evaluation criteria

for flexible business process techniques, both functional and non-functional are proposed. Then Section 4 practically demonstrates the way the evaluation criteria may be used during flexible PMS selection. Next, Section 5 discusses related work. Finally, Section 6 presents some concluding remarks and future work.

2 FLEXIBLE BUSINESS PROCESSES

Current business processes need to be flexible, in order to efficiently support the continuous changes that organizations undergo in their attempt to survive in today's volatile environment. These changes may be due to governmental regulation changes, changes of business goals and continuous innovation or due to changes in operational needs, such as improving performance, quality and generally optimizing business processes. In the rest of this section we briefly describe the types of business process changes that we are interested in and then we link them to available process flexibility approaches (Table 1).

There have been a lot of attempts to classify business process changes. Regev et al. (2006) suggest a change taxonomy based on three orthogonal dimensions: the abstraction level of change, the subject of change (e.g. organizational, operational, etc.) and the properties of change (e.g. extent, duration etc.). Leoni (2006) provides a hierarchical categorization of approaches supporting process adaptation based on the abstraction level of change, i.e. process type level and process instance level; approaches that deal with process instance changes are further categorized based on the kind of change they support, i.e. ad-hoc and preplanned changes; finally, approaches concentrating on pre-planned changes are classified based on the basic methods used for automatic failure detection and for change realization (e.g. goal-based, rule-based, etc.).

Our research interests are mostly concentrated on the abstraction level of change, the point in time that the change is taking place and its duration. Therefore, we provide a categorization of process changes based on two dimensions, while we show how they are inter-related (see Fig. 1):

a) The abstraction level of change, which may be, either at the process definition level or at the process instance level, also referred as evolutionary and ad hoc changes respectively (Rinderle, et al., 2004b). Changes at the type level are permanent and influence all process instances. Changes at the instance level affect only one case or a selected group of cases which means that it is not necessary to alter the business process definition.

b) A temporal dimension which considers both

i. The duration of change which may be permanent or temporary; alterations in permanent instance changes remain valid until the completion of the process instance while in temporary instance changes alterations may remain valid, for example, only until the completion of a one loop iteration of the current process instance.

ii. The timestamp of change occurrence which may be either at design-time (prior to process execution) or at run-time (during process execution). Design-time changes may be at entry-time (Mulyar, et al., 2007) (which can be either at instance level or at the process level, thus affecting only future instances). Run-time changes may be predefined changes (Regev et al., 2006) that occur at the process instance level, or on-the-fly changes (Regev et al., 2006) that may also be either at the instance level, which affect only the running instance or at the process type level, thus affecting both present and future process instances.

The implementation of the previously mentioned process changes should not be followed by the complete redesigning of the existing business process. A number of approaches have been designed to address this need and can be used in isolation or in combination. Namely, these approaches are Flexibility by Under-Specification, Flexibility by Deviation, Flexibility by Change and Flexibility by Design. An extensive description of these approaches may be found in (Schonenberg et al., 2008).

Each of these approaches addresses some of the changes that we have presented in Fig. 1.
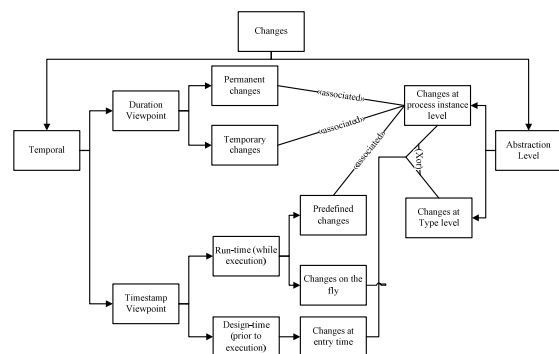
Figure 1: Change types and their relationships.

Table 1 depicts the types of change supported by each process flexibility approach. It is worthwhile noticing that *Flexibility by Under-Specification* and *Flexibility by Change* address most of the change types of our categorization.

References to the different *Change Types* and *Flexibility Approaches* are made in the next section where our proposed *Criteria Framework* for the selection of a flexible PMS is discussed.

Table 1: Change types and their association to process flexibility methods.

Approaches / Change Types	Flexibility by Design	Flexibility by Deviation	Flexibility by underspecification	Flexibility by Change
Change at Type level	√			√
Change at Instance level		√	√	√
Changes at entry time			√	√
Changes on the fly		√	√	√
Predefined changes			√	
Permanent instance change			√	
temporary instance change			√	

3 EVALUATION CRITERIA FOR FLEXIBLE BUSINESS PROCESS TECHNIQUES

In this section we identify a set of criteria that flexible PMSs should be evaluated against, derived both from literature study (Leoni, 2006; Pesic & van der Aalst, 2007; Rinderle et al., 2004a; Weber et al., 2007) and our experience while working with different such systems.

The suggested Criteria Framework may be used by anyone willing to select a flexible PMS amongst any number of such systems. We should note that a flexible PMS should not necessarily address equally all the functional criteria but rather focus on those associated with the specific types of changes and flexibility approaches that it supports. Thus, while we presenting each functional criterion, we specifically refer to the business process change types or business process flexibility approaches supported (as presented in Section 1).

The criteria are differentiated between functional, i.e. related with what the system provides, and non-functional, i.e. related with how the system performs, e.g. how secure it is or how easy it is to use.

3.1 Functional Criteria

- *Change Traceability.* It is important for all types of process change and can be utilized by all process flexibility approaches. The need for such mechanism may be driven by various reasons,

such as legal reasons, re-usability in case of a similar future change, conflict resolution and so on. Thus, this criterion needs to be supported by any flexible PMS.

- *Reuse.* It is needed in situations where the process definition is deviated very often, e.g. in the banking sector where customers with similar requirements may have to be serviced in a daily base. The reuse criterion should be primarily addressed by PMSs that support changes at the instance level and use the approaches of flexibility by change, by under specification and by deviation.

- *Change Concurrency Control.* In the today's volatile environment with highly-structured and long-lived business processes, different users/groups may need to implement simultaneously process changes, at the same process abstraction level or at different abstraction levels. Therefore, mechanisms for allowing changes in a controlled manner, avoiding severe errors and inconsistencies, are required for both process type and instance changes.

- *Migration Control.* It refers to the ability of a system to decide whether a change introduced in its process definition should affect a running process instance. This criterion is related to run-time, type changes and especially to changes on-the-fly and is required in PMSs using the flexibility by change method.

- *Version Control.* There are different ways to implement changes to a process definition which produce different variations of the initial process definition. Therefore, a versioning control mechanism that allows the co-existence of all the different versions, each tied together with its process instances is important for PMSs which use flexibility by design and change methods and support process type changes both prior and during process execution.

- *Change Impact Analysis.* The ability of a PMS to answer questions like "what is the impact of change?" is necessary in order to handle a large amount of candidate changes that may appear concurrently, examine if their implementation will be at the process instance level or at the overall process, prioritize them and even prohibit the occurrence of some of them.

- *Process Optimization.* It refers to the ability of a system to analyze process changes, focusing at the process instance level and then suggesting possible extensions/changes of the existing process definition. This criterion is mostly

applied to systems that provide flexibility by change.

- *Automation.* It refers to the ability of a system to provide automatic detection of process malfunction and automatic decision making on process improvements. It is useful to systems that support the process flexibility by change, by under-specification and by deviation approaches.

3.2 Non-functional Criteria

- *Specification Technique.* This affects the process flexibility of a system (Pesic & van der Aalst, 2006). Specifications based on imperative techniques (which describe how different tasks are linked) are not as flexible as those based declarative techniques that concentrate on the description of the different tasks that constitute a business process.
- *Correctness.* It refers to the absence of deadlock-causing cycles or erroneous data flows that may be triggered by changes. A way to achieve this is the existence of correctness criteria (Rinderle et al., 2004a) in order to check and ensure that only process instances compliant with the changed process schema are eligible to be updated.
- *Security.* It refers to the provision of privilege control mechanisms for process changes. Access rights should be simple to define in different levels of granularity and easy to maintain. Balance between flexibility and security is also important (Weber et al., 2004).
- *User-Friendliness.* It refers to the provision of adequate support, such as graphical interfaces, hiding technical details, so that change is facilitated by users. This criterion is important, regardless the process flexibility methods used by a system.
- *Response Time.* This criterion refers to the ability of a system to react to environmental or operational changes by deploying new processes in a timeless manner. Therefore appropriate measurements need to be provided in order to test the time of applying such changes.

4 SELECTING THE 'BEST-FIT' PMS

In this section we aiming at put in practice the proposed criteria of the previous section and follow

a set of guidelines that may assist a stakeholder, during the selection process of the 'Best-Fit' flexible PMS. The proposed guidelines are quite simplified in order to be easily followed by both business and technical oriented stakeholders and can provide an immediate, quantitative and accurate result. These guidelines are summarized next (Fig.2):

- Use the criteria framework of Section 3 and assign a weight to each evaluation criterion, based on its importance for a specific stakeholder.
- Evaluate candidate PMSs and give appropriate marks to each result, using a predefined metric system.
- Calculate scores for candidate PMSs and choose the one with the highest score.
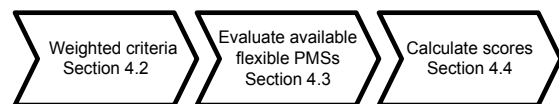
Figure 2: Selection process guidelines.

In the following section we will demonstrate how the proposed evaluation criteria and the aforementioned guidelines can be used by a Bank that needs to deploy a PMS to cater for its loan origination procedures. We have chosen to use the specific business case as we wish to take advantage of our expertise in this industry and share our accumulated experience.

4.1 Business Case Scenario Description

A Loan Origination System is part of the mission critical infrastructure of a typical retail banking organization; as financial products depended on such platforms contribute a large proportion to the operating margins of a commercial bank. In brief, a Loan Origination System (LOS), handles the steps taking place from the moment a customer applies for a loan product to the final approval (or not) of the request – including all decision logic –, its forwarding to the core banking system for the requested amount to be credited to the requestor and the archiving of the application along with any attachments (typically all documents needed for the approval of the loan) so it can be retrieved upon request.

4.2 Weighting Criteria

In order to weight each evaluation criterion, we decided to use a scale of 1-5; 1 indicates that the specific criterion has the lowest importance for the

bank and the business process at hand, whereas 5 indicates the highest importance. The provided weights appear in Table 2 followed by their justification.

Table 2: Weights of evaluation criteria.

Evaluation Criteria	Weight	Evaluation Criteria	Weight
Change Traceability	5	Automation	5
Reuse	4	Specification Technique	2
Change Concurrency Control	2	Correctness	5
Migration Control	1	Security	5
Version Control	5	User-Friendliness	4
Change Impact Analysis	5	Response Time	4
Process Optimization	4		

4.2.1 Justification for the Weights Provided for the Functional Criteria

- *Change Traceability.* Process change traceability is of significant importance for our case, since traceability serves a number of different purposes: comply with the regulatory framework imposed to all commercial banks, comply with internally set commercial policies, allocate financial incentives to employees involved in the approval process, etc. (Weight: 5)
- *Reuse.* The rapid growth of the financial services market, the intense competition characterizing the industry and the highly diversity in customer requests have established the capabilities: adaptation to all changing requirements and quick response to customer requests at process instance level, a key characteristic for financial organizations. Under the circumstances, only the reuse of previously used processes can guarantee such a response, ensuring at the same time low operational costs and quick response rates to customer requests, even if those requests involve non typical business scenarios. (Weight: 4).
- *Change Concurrency Control.* As banks operate in a heavily organizational structured environment, the case of numerous process changes taking place at the same time (either at the type level or at the instance level) is considered a low probably option. (Weight: 2)
- *Migration Control.* Migration control is not important in our business case as we need each instance to be tied to the process version that was active at the time it started its execution. This is because the business process definition has been communicated and agreed with the customer prior to the process instance execution. (Weight: 1)
- *Version Control.* It represents one of the most important criteria that need to be met by the PMS that is going to be selected for this case. Considering the long duration of the relationship

created between the bank and the customer when referring to loan products – this is especially true in the case of mortgages – the need to keep a highly efficient and trusted version control mechanism is of paramount importance. Any changes, either major or minor, must result to a new version, marked appropriately; also there is a need for co-existence of all the different versions each tied together with its process instances. (Weight: 5)

- *Change Impact Analysis.* A mechanism that checks whether a newly introduced change results to subsequent changes to the overall process itself, in order to avoid potential process inconsistencies, is very important when designing a loan origination system. (Weight: 5).
- *Process Optimization.* As any large organization is in constant need of optimized processes to ensure economically efficient output, this criterion is deemed quite important for our case at hand. (Weight: 4)
- *Automation.* One of the most important features of a PMS that deploys loan origination procedures, or at least an advanced one, is its ability to automatically adapt to new input and improve its processes. Most often those improvements can be triggered by the outcome of a process such as the number of loans that have been accepted or rejected. In addition, the automation of the decision making is very important for such a system. (Weight: 5)

4.2.2 Justification for the Weights Provided to the Non-functional Criteria

- *Specification Technique.* The definition of Loan Origination procedure is not as complex as other banking procedures like Asset Portfolio Management. For that reason both imperative and declarative specification techniques could be equally applied, minimizing the importance of this criterion. (Weight: 2)
- *Correctness.* It is a fundamental issue for any process management system. Especially when referring to a loan origination system of a bank, an error free operation is even more critical as it handles sensitive customer data (usually a lot of personal and financial information). Not to forget that errors discovered late may result to serious sanctions for the bank. (Weight: 5)
- *Security.* The security robustness and the access privileges, to the loan origination procedures of the bank, are very important. Access rights granularity is a related issue which must be

additionally addressed to enable the bank to "match" the LOS platform to the various roles and access rights found in the branches, business units, the call center, etc. (Weight: 5)

- *User-Friendliness.* The Loan Origination procedures are handled roughly by two types of users: the end users who are typically business users and the powers users (who are typically entitled to administrative rights). User-friendliness is important for the end users as they are not experienced users. On the other hand, user-friendliness is not of high importance for the power users as they are well trained technical people. However user-friendliness helps power users to ensure a rapid time-to-market in the accommodation of any new requirement of their business environment. (Weight: 4)

- *Response Time.* The volume of concurrent transactions by business users, along with the great amount of process instance changes that such a system may undergo in a major commercial bank, dictate the response time to changes as quite important. It should be noted that the response time of type changes is not considered, as they are implemented during off business hours. (Weight: 4).

4.3 Evaluation of Candidate PMSs

In this section we demonstrate how to use the *Criteria Framework* defined in Section 3, to evaluate PMSs that provide process flexibility. Also, we quantify the evaluation results by assigning marks to each one of them based on pair-wise subjective comparison.

4.3.1 PMS Evaluation Demonstration

For the specific case (bank) we chose to evaluate five flexible PMSs based on either their maturity or on their acceptance by the research community; these systems are: AristaFlow (Weber et al., 2004), YAWL (Dadam et al., 2007), FLOWER (Aalst & Berens, 2001), DECLARE (Pesic et al., 2007) and CAKE2 (Maximini et al., 2005). A summary with the results of our review can be found in Table 3.

It should be noted that the review was based on a comprehensive literature study, while actual tests were conducted for the systems that were available (eg. YAWL, DECLARE). Also the respective research groups were contacted for clarifications in some cases.

Table 3: Evaluation results of five process management systems.

Flexible PMS Criteria	AristaFlow	YAWL	FLOWer	DECLARE	CAKE
Change Traceability	fully satisfied: LogManager component (Weber et al., 2004)	partially satisfied: uses process logs (Workflow Logger) (Adams, 2007)	partially satisfied: uses Case guide (Aalst & Berens, 2001)	partially satisfied: uses execution logs	fully satisfied: uses execution logs enhanced with context (Maximini et al., 2005; Weber et al., 2007b)
Reuse	fully satisfied: ProcessRepository, ActivityRepository and ChangeOperations (Reichert et al., 2008)	fully satisfied: uses worklet repertoires	not satisfied	fully satisfied: uses constraint templates and recommendation service of PRoM (Process Mining Framework, 2008)	fully satisfied: uses WD characterization CB
Change Concurrency Control	partially satisfied: concurrent changes only across different process abstraction levels	fully satisfied: concurrent changes within the same process abstraction level and across different abstraction levels	not satisfied	not satisfied	partially satisfied: concurrent changes at the process instance level using breakpoints (Minor et al., 2008)
Migration Control	fully satisfied: compatible instances with the new process schema are migrated (ProcessManager component) (Weber et al., 2007b)	not satisfied	not satisfied	fully satisfied: process instances are migrated	not satisfied
Version Control	fully satisfied: ProcessRepository component (Reichert et al., 2008)	fully satisfied: uses Yawl engine and Process Repository (Adams, 2007)	partially satisfied: FLOWer studio (Aalst & Berens, 2001)	not satisfied	not satisfied
Change Impact Analysis	not satisfied	not satisfied	not satisfied	not satisfied	not satisfied
Process Optimization	partially satisfied: uses mining techniques (Günther et al., 2006)	not satisfied	not satisfied	not satisfied	not satisfied
Automation	not satisfied	not satisfied	not satisfied	not satisfied	not satisfied
Specification Technique	partially satisfied: uses imperative specification approach and adaptation patterns	partially satisfied: uses imperative specification approach (Yawl) with high expressiveness	partially satisfied: uses an imperative specification approach	fully satisfied: uses both imperative and declarative process specification techniques	partially satisfied: uses an imperative specification approach and weakly structured workflows
Correctness	fully satisfied: uses compliance criterion and pre/post-conditions (Reichert et al., 2008)	fully satisfied: uses YAWL process editor (Weber et al., 2007b)	partially satisfied: is ensured only for redo/skip actions of process instance level	fully satisfied: checks for constraint violation	fully satisfied: using a suspension mechanism (Minor et al., 2008)
Security	fully satisfied: uses OrgModelManager (Reichert et al., 2008)	fully satisfied: uses access rights	fully satisfied: uses working profiles (Aalst & Berens, 2001)	not satisfied	partially satisfied: uses a login mechanism
User-Friendliness	not satisfied	partially satisfied: the cornerstone case is automatically presented to the user	not satisfied	fully satisfied: uses constraint templates, recommendation service of ProM, integration with YAWL	partially satisfied: CAKE editor (Maximini et al., 2005)
Response Time	not satisfied	not satisfied	not satisfied	not satisfied	not satisfied

The evaluation results provide an insight into the manner and extent to which the criteria are satisfied by the selected PMSs. Important evaluation remarks follow:

- AristaFlow scores well to all criteria, but lack support for user-friendliness, automation and response time. It assumes that changes are performed manually, by expert users.

- AristaFlow and CAKE provide a more complete mechanism for change traceability as their process change logs are enriched with contextual information, related to the reasons for those changes.

- YAWL and AristaFlow enable concurrent changes across different process abstraction levels as a new process version is not necessarily followed by instance migration. Also both

CAKE and YAWL allow concurrent changes at the process instance level. In YAWL changes are restricted to local placeholder activities, meaning that different placeholder activities can be concurrently modified. On the other hand in CAKE only the parts that have to be modified are suspended (using breakpoints) during change, while parallel branches, not affected by that change, can proceed with their execution (Minor et al., 2008).

- FLOWer is widely adapted by organizations. However in our evaluation it seems to be the weakest system, as most of the criteria are not met. These results may be explained if we consider the systems' rationale which is based on the case handling paradigm and supports run-time process deviations. Thus, for instance, the existence of a reuse mechanism is not as vital as it is for the other systems that support run-time changes and use different flexibility approaches. Also, its versioning mechanism does not provide an accurate solution (Weber et al., 2007b).

- DECLARE stands out for its specification technique and its user-friendliness. It mostly uses a constraint-based declarative language. However, complex business processes are specified using an imperative specification technique. Besides, user-friendliness is reinforced by a graphical notation for constraint templates. Users are also assisted while deciding on the order of tasks, by the recommendation service of ProM which compares the current process instance with past executions and favors those executions that satisfy the specified goal.

- AristaFlow is the only system that provides adequate mining techniques (Günther et al., 2006) to the change log files of modified process instances of the ProcessRepository. The results of such analysis may be used for future process improvements (Dadam et al., 2007).

- The evaluated systems, except DECLARE, use an imperative process specification technique. However YAWL and CAKE are the only ones that support weakly structured process definition using late planning and late modeling and late binding (Maximini et al., 2005).

4.3.2 Quantifying Evaluation Results

In order to measure the appropriateness of each PMS we define a metric system by assigning marks - using a scale from 1 to 3 - to the evaluation results of Table 3. Thus, when a criterion is fully satisfied, it is marked with '3', partially satisfied it is marked with '2', and not satisfied it is marked with '1'. The results, for each PMS, are depicted in Table 4, under the columns titled as 'Mark'.

4.4 The 'Best-Fit' PMS

Finally, using the marks of the evaluation results and by combining them with the weights originated from Table 2, we are able to quantify the appropriateness of each process management system to provide support to the specific stakeholder for its loan origination procedures. The results are summarized in Table 4, under the column 'Score'.

Table 4: Measurement of the appropriateness of each PMS for the Loan Origination procedures.

Flexible PMS / Criteria	AristaFlow		YAWL		FLOWer		DECLARE		CAKE	
	Mark	Score	Mark	Score	Mark	Score	Mark	Score	Mark	Score
Change Traceability	3	15	2	10	2	10	2	10	3	15
Reuse	3	12	3	12	1	4	3	12	3	12
Change Concurrency Control	2	4	3	6	1	2	1	2	2	4
Migration Control	3	3	1	1	1	1	3	3	1	1
Version Control	3	15	3	15	2	10	1	5	1	5
Change Impact Analysis	1	5	1	5	1	5	1	5	1	5
Process Optimization	2	8	1	4	1	4	1	4	1	4
Automation	1	5	1	5	1	5	1	5	1	5
Specification Technique	2	4	2	4	2	4	3	6	2	4
Correctness	3	15	3	15	2	10	3	15	3	15
Security	3	15	3	15	3	15	1	5	2	10
User-friendliness	1	4	2	8	1	4	3	12	2	8
Response Time	1	4	1	4	1	4	1	4	1	4
Final Score		109		104		78		88		92

We should note that the scores associated with each criterion, for every PMS (Table 4), have been derived using the following formula: (Criteria Weight) x (Criteria Mark) = (Criteria Score). Conclusively, the system with the highest final total score is AristaFlow, which is clearly the one that should be selected by the Bank.

5 RELATED WORK

There are a few publications dealing with the establishment of evaluation criteria for comparing flexible Process Management Systems. For instance Selmin Nurcan (2008) has introduced such criteria based on properties like: nature of flexibility, nature of impact, etc. However this work does not demonstrate how the criteria may be applied in practice by evaluating flexible process management systems. Also Helen Schonenberg (2008) proposes an extensive taxonomy of process flexibility that is used to evaluate a PMS systems. This taxonomy is focused on process flexibility approaches and their characteristics (eg. deviation/change operations, migration strategies for evolutionary changes, etc.). Correctness criteria are analyzed by Stefanie Rinderle (2004a) and are used, along with modeling properties, to evaluate approaches supporting flexible workflows like WIDE, Breeze, etc. Change

patterns and change support features are introduced and used by Barbara Weber (2007a) to assess the power of process change frameworks like WIDE, MOVE, HOON, etc. The change patterns include both adaptation patterns and patterns for predefined changes. On the other hand change support features include: schema evolution, instance migration, support for ad-hoc changes, correctness of change, traceability analysis, etc. Our criteria framework extends the work of Barbara Weber (2007a) by incorporating concepts like change concurrency control, change impact analysis, specification techniques, process optimization, use-friendliness, change response time and change automation.

Besides, there is research work dealing with the evaluation of business process management systems. A paper that analyses the state-of-the-art of such evaluation efforts is published by Andreas Schmietendorf (2008). It provides an analysis of available evaluation approaches especially for business process modeling tools and produces an empirical evaluation of Business Process Management tools based on criteria like supporting modeling notations, interface formats, report functionalities, degree of relevance, etc.

Apart from these approaches that mostly introduce criteria and evaluate different PMSs, we are not aware of any work that defines a *Criteria Framework for flexible PMSs* and provides guidelines on how to use such criteria-based approach to select from a set of admissible flexible PMSs, while demonstrating them practically and sharing experiences using a specific case coming from a major industry (e.g. Banking).

6 CONCLUSION AND FUTURE WORK

In this paper we stressed the need for flexible business processes. Also we identified a *criteria framework* that flexible PMSs should comply with. Finally we demonstrated the way that the criteria framework may be applied practically during the selection of the 'best-fit' flexible PMS, using a realistic business case originating from the banking sector.

Finally, future work could focus on: (1) using the acquired knowledge by applying the criteria framework to evaluate existing flexible PMSs and provide appropriate mechanisms for supporting criteria that were not satisfied at all, or were partially satisfied (e.g. change impact analysis, optimization, automation, etc.) by such systems, and (2) applying

the *Criteria Framework* and related guidelines to other major industries.

REFERENCES

Aalst, W. M. P., Aldred, L., Dumas, M., and Hofstede, A.H.M., 2004. Design and Implementation of the YAWL System. *In Proceedings of the 16th International Conference on Advanced Information Systems Engineering (CAiSE'04)*, volume 3084 of Lecture Notes in Computer Science, pages 142–159. Springer Verlag.

Aalst, W. M. P., Berens, P.J.S., 2001. Beyond Workflow Management: Product-Driven Case Handling. In S. Ellis, T. Rodden, and I. Zigurs, editors, *International ACM SIGGROUP Conference on Supporting Group Work (GROUP 2001),* pages 42–51. ACM Press, New York.

Adams, M., 2007. *Facilitating Dynamic Flexibility and Exception Handling for Workflows*. Phd thesis. Faculty of Information Technology, Queensland University of Technology, Brisbane, Australia (2007)

Dadam, P., Reichert, M., Rinderle, S., Jurisch, M., Acker, H., Göser, K., Kreher, U., Lauer, M., 2007. "ADEPT2 - Next Generation Process Management Technology". *Heidelberger Innovations forum*, Heidelberg.

Günther, C.W., Rinderle, S., Reichert, M., van der Aalst, W.M.P., 2006. Change Mining in Adaptive Process Management Systems. Proc. *14th Int'l Conf. on Coop. Information Systems (CoopIS'06)*, Montpellier, France, Nov. 2006, LNCS 4275, pp. 309-326.

Leoni, M., 2006. Adaptive Process Management, 20 October 2006.

Maximini, R., Freßmann, A., Sauer, T., Maximini, K., Bergmann, R., 2005. CAKE - Collaborative Agent-based Knowledge Engine. Technical Report, University of Trier, Department of Business Information Systems II.

Minor, M., Tartakovski, A., Schmalen, D., Bergmann, R., 2008. Agile Workflow Technology and Case-Based Change Reuse for Long-Term Processes. International Journal of Intelligent Information Technologies, 4(1):80-98.

Mulyar, N.A., Schonenberg, M.H., Mans, R.S., Russell, N.C., Aalst, van der W.M.P, 2007. Towards a taxonomy of process flexibility (extended version). Technical report.

Nurcan, S., 2008 "A Survey on the Flexibility Requirements Related to Business Processes and Modeling Artifacts," in Proceedings of the 41st Hawaii International Conference on System Sciences HICSS'2008, pp. 378-378.

Pesic, M., van der Aalst, W.M.P. 2006. A declarative approach for flexible business processes management. In Business Process Management Workshops, pages 169–180.

Pesic, M., Schonenberg, H., van der Aalst, W.M.P., 2007. DECLARE: Full Support for Loosely-Structured

Processes. In M. Spies and M.B. Blake, editors, Proceedings of the Eleventh IEEE International Enterprise Distributed Object Computing Conference (EDOC 2007), pages 287–298. IEEE Computer Society 2007.

Process Mining Framework, 2008. http://prom.win.tue.nl/tools/prom/

Regev, G., Soffer, P., Schmidt, R., 2006. Taxonomy of flexibility in business processes. In Proceedings of the 7th Workshop on Business Process Modelling, Development and Support (BPMDS'06), 2006. http://lamswww.epfl.ch/conference/bpmds06/taxbpflex.

Reichert, M., Dadam, P., Juritsch, M., Kreher, U., Göser, K., Lauer, M., 2008. Architectural design of flexible process management technology. Technical Report 2008-02, Univ. of Ulm, 2008.

Rinderle, S., Reichert, M., Dadam, P., 2004a. Correctness criteria for dynamic changes in workflow systems - a survey. *Data and Knowledge Engineering, Special Issue on Advances in Business Process Management*, 50(1):9-34, 2004.

Rinderle, S., Reichert, M., Dadam, P., 2004b. On Dealing with Structural Conflicts between Process Type and Instance Changes. In *Second International Conference on Business Process Management*, pages 274{289, Postdam, Germany, June 2004.

Schmietendorf, A., 2008. *Assessment of Business Process Modeling Tools under Consideration of Business Process Management Activities*, in Lecture Notes in Computer Science, Volume 5338/2008, pp. 150–163, Springer-Verlag Berlin Heidelberg 2008

Schonenberg, H., Mans, R., Russell, N., Mulyar, N., van der Aalst, W.M.P, 2008. Process flexibility: A survey of contemporary approaches. In *Proceedings of Advances in Enterprise Engineering I, 4th International Workshop CIAO! and 4th International Workshop EOMAS, held at CAiSE 2008*, Montpellier, France, June 16-17, 2008.

Weber, B., Rinderle, S.B., Reichert, M.U., 2007a. Change patterns and change support features in process-aware information systems. In Proceedings *19th International Conference on Advanced Information Systems 212 Engineering (CAiSE 2007)*, volume 4495 of Lecture Notes in Computer Science, pages 574–588, Trondheim, Norway, June 2007. Springer Verlag. 175, 176, 178

Weber, B., Rinderle, S.B., Reichert, M.U., 2007b. Change support in process-aware information systems - a pattern-based analysis. *Technical Report Technical Report TR-CTIT-07-76, ISSN 1381-3625*, Centre for Telematics and Information Technology, University of Twente, Enschede, 2007.

http://eprints.eemcs.utwente.nl/11331/

Weber, B., Reichert, M., Wild, W., Rinderle, S., 2005. Balancing flexibility and security in adaptive process management systems. In R. Meersman and Z. Tari, editors, Proceedings of the *13th International Conference on Cooperative Information Systems (CoopIS'05)*, volume 3760 of Lecture Notes in

Computer Science, pages 59–76, Agia Napa, Cyprus, November 2005. Springer Verlag. 35

Weber, B., Wild, W., Breu, R., 2004. CBRFlow: Enabling adaptive workflow management through conversational case-based reasoning. In Peter Funk and Pedro A. Gonzalez Calero, editors, *Proceedings of the 7th European Conference for Advances in Case Based Reasoning (ECCBR'04)*, volume 3155 of Lecture Notes In Computer Science, pages 434–448, Madrid, Spain, August 2004. Springer. 22.

BUSINESS TERMS
Model for a Telecom Operator Business View of SLA

Carlos Bueno Royo, Juan Lambea Rueda, Óscar L. Dueñas Rugnon, Beatriz Fuentes
and Alfonso Castro

Telefónica Investigación y Desarrollo (Telefónica I+D), Emilio Vargas 6, Madrid, Spain
{carlosb, juanlr, oldr, fuentes, acast}@tid.es

Abstract: The rapidly growing service-oriented economy has highlighted key challenges and opportunities in ICT-supported service and product lifecycle management. From a service consumer point of view there is no standardized way to locate, evaluate, negotiate and monitor services. Composition of third party services is fraught with uncertainty due to absence of definitive SLAs. Moreover, there is a lack of a standardized, homogenized and extended-used model for a set of business terms for SLAs that allows an automatic negotiation of products with customers and with third parties. So, from the service provider perspective, creating customized service offerings, negotiating with individual customers, and translating from business requirements into specific internal provisioning manifestations consumes valuable time and resources. Furthermore, there is no mean to create attractive products composed by several different services coming from different domains and consolidate the underlying SLAs in a business customer-faced SLA. Ultimately, the service marketplace is frustrating and cumbersome for both service providers and consumers: a significant opportunity has arisen for a holistic SLA-management framework, able to understand and manage such a set of integrated and homogenized set of business parameters.

1 INTRODUCTION

The ongoing transformation of a product-oriented economy towards a service-oriented economy has come to a critical point. IT-supported service provisioning, aggregation and commercialization have become of major relevance in all industries and domains. However, the nature of these services is typically quite static, because it requires significant effort to create service offers, to negotiate provisioning and business details with customers and to manage and control provided services.

From a business point of view, there is a need to hide provisioning low level details to customers. In this way, there is a gap in the management of Business SLAs that drives the product lifecycle. Business SLAs, which contain basically business terms for the implied services in the product sale, have to be translated to lower layers SLAs (software and infrastructure levels), so the product provision could be properly performed (Theilmann,2008). From this perspective, it is essential to provide a mean to create true integrated Business SLAs from atomic services and build product offerings with a

homogeneous set of business terms that can be rapidly understood by customers, hiding lower level details.

Furthermore, in multi-provider and multi-domain scenarios where is possible to create products with services coming from different domains, it is essential to be able to consolidate the different SLAs of the composed service into a unique business SLA. Moreover, the set of business terms proposed has to be used between providers, so the composition of products from third-provider services can be easily and faster made. The integrated SLA will be the one negotiated with final customers, so that underlying services requirements, multi-domain provisioning difficulties and business relationship between providers are hidden to customers. The creation and composition of such a holistic SLA, and its following monitoring and assessment, is a challenge that must be faced in order to be able to build attractive and proper product offers in the top of the future frameworks and clouds (Comuzzi, 2009).

The aim of this paper is try to make a definition of the needed business terms for B2B transactions that have to be consumed by final customers. New business environments need to be automated in some

issues, such as product discovery and contracting. But it is also essential to define the business terms to be agreed. Furthermore these relationships between customers and providers have to be supported by business SLAs built from artefacts, terms and features that are introduced in this paper, such as business terms.

To do so, the paper firstly briefly summarizes the state of the art on the industrial approaches related to business SLA models. Following, it is defined and explained a business model. In section 4, the paper proposes a set of business terms in which the SLA model could be based. Next, it is presented how the business model could be applied in the telecom industry. Finally, we state the main conclusions extracted from the paper and the future works in this line.

2 INDUSTRY APPROACHES

There have been many efforts to try to consolidate a set of definitions concerning enterprise information, especially business data. Therefore, there are many standardization organizations working on this; the TMForum (TeleManagement Forum, 2010) and the ITSMF (Information Technologies Service Management Framework) (ITSMF, 2010) currently are the most relevant at ICT business level. Consequently, the most important ones in this field have been achieved in initiatives, such as: SID (Shared Information and Data Model) (TeleManagement Forum-SID, 2010), ITIL (Information Technology Infrastructure Library) (ITSMF-ITIL, 2010) and eTOM (enhanced Telecom Operations Map) (TeleManagement Forum-eTOM, 2010).

The eTOM is the most widely used and accepted standard for business processes in the telecommunications industry. The eTOM model describes the full scope of business processes required by a service provider and defines key elements and how they interact. This model also includes a specific process for SLAs.

On the other hand, ITIL concentrates on IT Service Management capabilities, focused on the kind of services and service management that IT systems provide within organizations, independent of the kind of business or industry sector concerned. These reference frameworks (eTOM and ITIL) can be considered as complementary. eTOM provides the "big picture" while ITIL details the operations processes addressed and a view of process interfaces.

The other initiative identified is the SID, but this will be explained in detail in next subsection. However, the relationship between eTOM and SID is following the NGOSS initiative (TeleManagement Forum-NGOSS, 2010), where first, it is necessary to establish a business analysis, and then, it is defined the data analysis. In the NGOSS initiative eTOM supports the business analysis and SID initiative supports the data analysis.

2.1 SID Model

SID is a set of comprehensive standardized information definitions aiming to be the common language for building easy to integrate OSS (Operational Support System) and BSS (Business Support System) solutions. It has a specific model for SLA agreements.

The SID model focuses on what are called "business entity" definitions and associated attribute definitions. A business entity is an actor or object of interest to business such as customer, product, service, or network, with attributes describing the entity. Together, the definitions provide a business-oriented perspective of the information and data that it is needed to run in an organization.

With eTOM, the SID model provides enterprises with not only a process view of their business but also an entity view. That is to say, the SID provides the definition of the 'things' that are to be affected by the business processes defined in the eTOM. The SID and eTOM in combination offer a way to explain 'how' things are intended to fit together to meet a given business need.

The SID model also defines attributes for Business entities and relationships between them and is represented using an UML model that provides an architecturally oriented business view of business entities, their attributes, and relationships to other business entities.

In the following figure, the domains and the level 1 of Agreements Business Entities of the SID Business Model are presented:
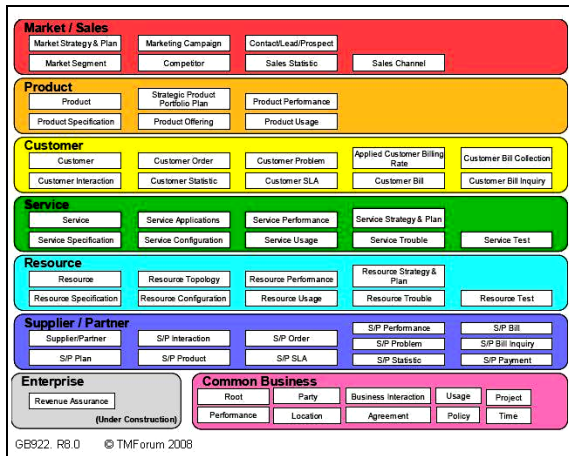
Figure 1: SID domains.

SID model has a common entity specific for agreements called Agreement. An agreement is a contract or arrangement, either written or verbal and sometimes enforceable by law, which involves a number of other business entities, such as Services, and/or Resources.

SLA is a type of agreement represented in the SID model by the Service Level Agreement entity. Both entities – Agreement and Service Level Agreement – consist of items which make the link between the SLA and the Product/Service/Resource it applies to.

For this purpose, SID Service Level Agreement Item entity establishes relations with entities in different domains.

2.2 SLA Management Handbook

There is another relevant reference framework, the SLA Management Handbook series (TeleManagement Forum-SLA, 2010), this is a handbook to assist two parties in developing a Service Level Agreement (SLA) by providing a practical view of the fundamental issues. The parties may be an "end" Customer, i.e., an Enterprise, and a Service Provider (SP) or two Service Providers. In the latter case one Service Provider acts as a Customer buying services from the other Service Provider. For example, one provider may supply network operations services to the provider that supplies leased line services to its customers. These relationships are described as the Customer-SP interface and the SP-SP interface.

The perspective of the SLA Management Handbook series is that the end Customer, i.e., an Enterprise, develops its telecommunication service requirements based on its Business Applications.

These requirements are presented to a Service Provider and the two parties begin negotiating the specific set of SLA parameters and parameter values that best serves both parties. For the SP, the agreed-upon SLA requirements flow down through its organization and become the basis for its internal management and control of its Quality of Service (QoS) processes. For the Enterprise Customers, the SLA requirements serve as a foundation or a component of its internal network services or business services. The SLA requirements define and limit the offer and the consumption of the service for both parties that are joined. It means that affect to the consumption of the customer and the offer, service level and support of the service by the service provider.

3 BUSINESS SLA MODEL

This section presents the Business SLA Model as a series of inter-related UML Class Diagrams. The starting point was the identification of the main concepts of the SLA model from the business point of view as basis for the construction of a Business SLA Model. The business models is then translated to a collection of Business SLA Parameters , which has been defined as business terms

This work is based on a generic SLA model defined in the SLA@SOI project and described in (SLA@SOI, 2009). Class diagram in figure 2, taken from (SLA@SOI, 2009) captures the main notions concerning SLAs. The class Agreement is a symbolic artefact representing a potential Agreement between a Customer and a Provider concerning the delivery of a Service – e.g. an agreement which has been proposed, but which may still be subject to further negotiation. The subclass SLA, instead, denotes an actual ServiceAgreement – i.e. one that has been formally accepted by both Customer & Provider.

Figure 2: SLA Model classes.

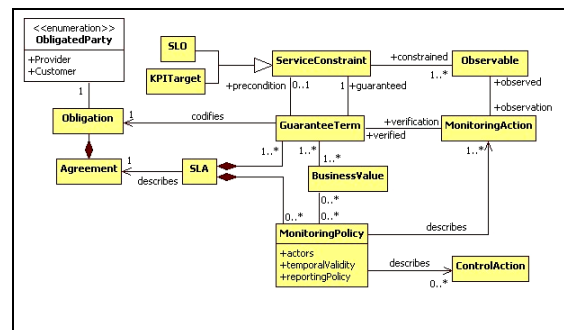The SLA comprises a set of GuaranteeTerms – each of which describes an Obligation on the part of the either the Customer or Provider to respect various ServiceConstraints. These Obligations, in turn, may be associated with various penalties or rewards - termed BusinessValues.

In the previous paragraphs we have described a generic SLA Model, valid to express the agreement conditions in any of the layers of the IT stack. In the following, this model will be refined introducing business-specific concepts.

This generic model can be particularized introducing business-specific concepts as service upgrade procedure, SLA update procedures, backup & recovery mechanisms, information about the service support system, termination information, tarification and billing .

The SLA may also contain appropriate monitoring policies – i.e. rules governing how monitoring of the service is going to be implemented. The SLA Monitoring model is further decomposed in Figure 3:
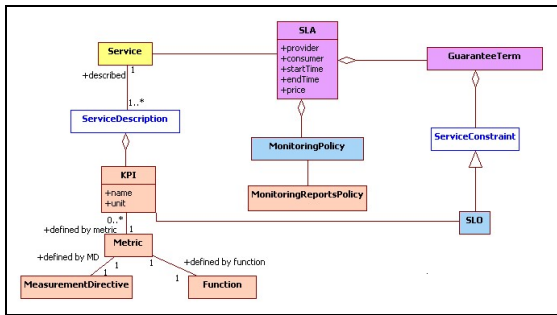
Figure 3: Monitoring model.

A KPI (Key Performance Indicator) is used to define measurable and exposed properties associated with a service, such as response time and throughput. The properties are used to express service level objectives (SLOs). The exact definition of a KPI is provided by a Metric, which also specifies how to measure or to compute the Performance Indicator.

A Metric is composed by a set of parameters that are measured from a service providing system (MeasurementDirective) or computed from other metrics and constants (Function).

The SLA has attached a set of MonitoringPolicy, which describes the monitoring activities: parameters to be collected and measures to be made, data collection and measurement intervals, and aggregation intervals.

The SLA should include a description of the monitoring reporting activities. , including reporting specification (content, format, frequency, delivery

mechanism...) and the description of the procedure to request changes in the reporting process.

Apart from the monitoring, one of the most important topics from the business point of view is the tariff and billing procedures. In Figure 4, the billing part of the SLA model has been magnified:
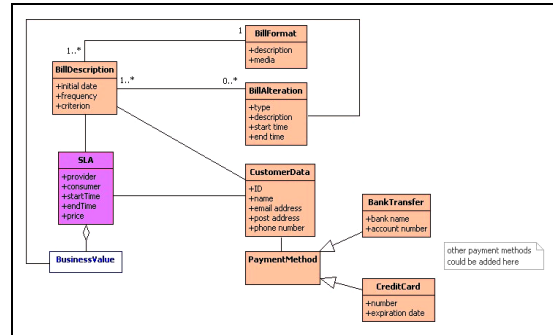
Figure 4: Billing concepts and model.

A BillDescription describes the detailed structure of the customer‟s bill. It indicates when to initiate the bill creation, as well as the frequency of the billing. It also specifies the billing mechanisms, detailing how the service is charged (i.e. for service usage).

This description should include a BillFormat, representing the presentation format of the bill. This bill may be sent to the customer through various media (for example, email, post mail ...).

The data needed for the billing process are represented by the Customer and PaymentMethod classes. Customer encapsulates all the customer data needed to contact him, i.e. name, address, e-mail, phone number. The PaymentMethod class represent the method used for payment. As an example, two subclasses have been included in the figure: BankTransfer, which includes the name of the financial institution and the account number where the bill will be charged, and the CreditCard, defined by the card number and the expiration date. Other payment methods can be added by extending the PaymentMethod class.

In some cases, the usual billing charge may suffer an alteration, represented by the BillAlteration class. The type attribute indicates the nature of the discount, i.e. a promotion, a reduction due to a violation of the signed SLA, or even an increase derived from a penalty applied to the Customer.

Next section shows in detail how this Business SLA Model has been translated into specific Business Terms to be included in the SLA document.

4 BUSINESS TERMS

In the following lines we explain the different aspects, terms and features that may be expressed at business level in the SLAs.

Functional Description: The offered service must be detailed in terms of the features and functionality supported and available to customers.

This information must to be completely enriched with service categorization and/or specification in order to be able to classified it. Also it must be appropriate to add semantics details that can be used to explain the full description of the service.

Business Model Supported: All the information referring to the selling process must be defined and described in detail, to avoid ambiguity for the customer. The business model description should detail:

- Offer types associated with the Quality of Service supported for each one
- Pricing model
- Billing and payment constraints
- Modification and alteration of prices if applicable
- Restrictions and constraints

Penalties: Detailed specifications about the penalties incurred when problems arise in the consumption of the service. This information is attached to the guarantee terms definitions that explain in detail how the different agreed terms are used. Each Guarantee Terms defined in the SLA must to be expressed in terms of parameters that express the conditions to be evaluated, the KPI target to be fulfilled and the evaluation characteristics to be taken into account for checking the generation of the penalty. In this way it is necessary to specify the periodicity of the checking explained before (assessment time interval) and the economic value associated to it to be generated.

Termination Clauses: The termination clauses have to be automated and they have to accommodate both parties in the contract. The termination of the SLA can be triggered by certain customer aspects as well as by certain service provider constraints. Also it must be defined the conditions below to be communicated to the parties.

Service Information Events, Monitorization and Reports: It must be possible for the final customer to select the kind of information that they wish to obtain automatically and how it has to be monitored. This information is defined in terms of events monitored as well as reports associated with the customer's service. For instance a customer that uses

a storage service may want to know how large their storage consumption is per day). This information is just used to report customer, and might not be correlated with Guarantee terms of the SLA. Features about the reporting that can be defined can be based on different parameters:

- Type of report (based on service activities parameters, only on penalties of the SLA, …)
- Communication channel for sending the reports to the customer (i.e. by email, by SMS/MMS,…)
- Report frequency (i.e. hourly, diary, monthly,…)
- Content customization of the report

Support Mechanism and Contact Details: It must be possible to specify the kind of support offered to the customer should they have a problem or inquiry. The support information provided should include timetable details as well as details of the different support channels available. For instance, it may specify a telephone number and the days and hours of availability. This information is complementary to the contact data of the customers or service providers such as, their whole names, telephone and fax numbers or the email address.

Other important aspects to be taken into account are the supporting procedures. It specifies how the support team will responds to an error communicated by the customer, that is, the level of severity of the errors and the associated times to respond and solve to them. The supporting procedures, usually define different levels of the support, based on the kind of problem that can arise. With the classification of the problem, the supporting procedures and timeline defined is different. That information is very important in the consumption of the services and it has to be clearly defined.

Disaster Recovery and Data Security in IT Systems: It must be possible to define Backup/Restore policies in order to guarantee the persistence of information, if the service offered to the customer manages and stores data. Also it must be possible to define the security mechanisms that are employed by the service. The terms and features needed for whole business service must to be concreted also in the SLA:

- Backup: Specifies how the system will make a backup of the important information in terms of the frequency, the mechanism type (full, mirror, incremental) and the estimated time needed to do this.
- Recovery: Specifies how the system will perform a service recovery in case of a fatal error in terms of the mechanism type (full, mirror, incremental) and the estimated time needed to do this.

Changes to Terms in the Service: The process to update the service conditions or characteristics must also be considered. It must also be possible to define the mechanism used to inform the customer about such changes. Also it must be possible to define the frequency of updating process, how the customer will receive the actualization, notification or the instructions to install it (if it is something that depends on the customer operation).

Customer/Provider Requirements and Constraints: In many cases, it is necessary for the customers or providers to express some requirements in terms of limits or constraints in the service consumption. Usually these aspects are related to legal constraints to be followed by the customers or providers of one specific country, because they are imposed by the relevant Regulatory Authority.

Example constraints may include:
- Personal data storage cannot to be stored outside the country
- Maximum prices and/or quality shall apply
- Restrictions in sharing of personal data with third parties associated with the service provider
- The prohibition of delivery advertisement
- Personal data usage restrictions for specific tasks (e.g. data mining)

5 BUSINESS APPLICABILITY

So far, we have showed a business SLA model and a set of terms associated to it. But how can they be applied in the Telecom operator business?

Until recent times, the classic business Telecom operator was associated with the network, even when its prior service, voice, was offered over an ad hoc one, the public switched telephone network (PSTN) that evolved from an analogue and fixed network to digital and mobile ones.

Nowadays the environment has changed. The advances in technology have allowed the modification of the service concept. So the service provider can offer services available at any time or place, accessible using any kind of access network and regardless of the kind of device you use. Also new roles have appeared: services providers, infrastructure providers, content provider, telco provider … as well as new actors which can assume different roles.

Google is a clear example: in its beginning Google was a simple information searching engine and nowadays it has announced its entry in the telecommunications business with its project to deploy a high-speed network in United States.

On the other hand, telecommunication companies are attempting to expand their traditional business and become smart players in the content management market. The boom of the mobility and the connected devices has open new business possibilities. The future trend is that mobile devices will be used for the remote management of services in our daily life as part of what has been called Digital Personal Environment. The content and the communication channels and devices will be deeply linked. Telcos are aware of the increasing importance of software, that will allow the management of this foreseen on-line world. Moreover, users become prosumers, participating in the service creation and generating income. Telecommunication companies want to take advantage of this, and they are already creating platforms to help the developers sharing with them the profit of the applications. Wholesale Applications Community is an example of this type of platforms. This alliance, being launched by twenty-four of the world largest phone operators, aims to make it easier for developers to build and sell apps irrespective of device or technology.

The challenge of entering new markets also implies that the business model of the Telecommunication companies must evolve from the traditional model of customer proximity, high margins and high commitment to a one with distant relationship with the customer, lower margins and lower commitment.

This new marketplace must support the registration of multiple heterogeneous services, from different providers (prosumers) and industries, guaranteeing at the same time the quality of the delivered services, and maximizing revenues. New relationships and agreements among participants will be established, allowing the creation of new services by composing atomic services of different providers (Bueno, 2008).

The new situation requires the coherent management of the supplied services, and Service Level Agreements (SLAs) are a mean for doing this. SLAs specify the conditions under which services are delivered, but they can also be used for managing the provisioning and runtime functioning of the services. The end-to-end governance of the service lifecycle expands across the business, software, infrastructure and network layers, and therefore the SLAs must also be translated across the different layers. In this environment a precise definition of the business terms comes of crucial

importance, since they subsume the relationship not only with the customer but also with a third-party provider. A Business SLA may specify for instance the levels of availability, quality of service, performance, billing methods and even the penalties when the agreement is violated. These business terms will allow the automatic negotiation with customers and third parties in this new marketplace, and the translation to lower level SLAs helps to the automation of the provisioning procedures and runtime QoS monitoring and adjustment.

6 CONCLUSIONS AND FURTHER WORKS

The Business SLA terms have been identified to fulfill and define the business relationship between customer and provider in automated systems that can be used in business to business approaches.

This approach is based on a deeply study of the Business SLA Model and other different works based on SID and eTom framework from TMForum.

The results of this paper identify and define a proper Business SLA model for telco industry approach. In this line, it is presented a set of business parameters that should be present in this kind of business SLA model and could help in the automation of SLA lifecycle processes, specially in negotiation processes between customers and providers. Moreaover, they are strongly recommended to be used in B2B environements in scenarios where customers consume multi-provider and multi-domain products.

Further works will be focused on the concretes parameters and values of the business terms in each specific aspect and also in the study of new marketplaces in which providers can co-create and reuse services from others.

This new products will be consumed by end customers who will not be aware of the services (and the providers and domains involved in those services) that may exist behind the consumed product. However, the added value services that are aggregated or composed from different parties, have to guarantee the quality of service offered and have to share the revenues and penalties that can be triggered under different circumstances of service consumption. Consequently and for obvious reasons, end customers will not be aware of this dependency and complexiy of the services.

We have to provide the novelty tools to build the new business environments than can interact in different marketplaces and scenarios.

ACKNOWLEDGEMENTS

REFERENCES

Bueno, C., González, F., 2008. E-marketplaces: A Telecom Operator Perspective, in: IAWTIC08, 2008 International Conferences on Intelligent Agents, Web Technologies and Internet Commerce. IEEE Computer Society

Comuzzi, M., Theilmann, W., Zacco, G., Rathfelder, G., Kotsokalis, C., Winkler, U., et al. 2009. A Framework for Multi-level SLA Management. ICSOC09

ITSMF, 2010. "Information Technology Service Management Forum" (Online). Available: http://www.itsmfi.org/, Retrieved: Feb 2010.

ITSMF-ITIL, 2010. "Information Technology Infrastructure Library" (Online). Available: http://www.itil-officialsite.com/home/home.asp, Retrieved: Feb 2010.

SLA@SOI project, A5 WP. 2009. D.A5a-M12-SLA Foundations and Management. http://sla-at-soi.eu/wp-content/uploads/2009/10/D.A5a-M12-SLA-Foundations-and-Management.pdf

Theilmann, W., Yahyapour, R., Butler, J., et al. 2008. Multi-level SLA Management for Service-Oriented Infrastructures, ServiceWave 2008.

TeleManagement Forum-SID, 2010. "Information Framework (SID)" (Online). Available: http://www.tmforum.org/InformationFramework/1684/home.html, Retrieved: Feb 2010.

TeleManagement Forum-eTOM, 2010. "Business Process Framework (eTOM)" Online. Available: http://www.tmforum.org/BusinessProcessFramework/1647/home.html, Retrieved: Feb 2010.

TeleManagement Forum, 2010. "TeleManagement Forum" Online. Available: http://www.tmforum.org/browse.aspx, Retrieved: Feb 2010.

TeleManagement Forum-NGOSS, 2010. "Solution Frameworks (NGOSS)" (Online). Available: http://www.tmforum.org/SolutionFrameworks/1911/home.html, Retrieved: Feb 2010.

TeleManagement Forum-SLA, 2010. "SLA Management" (Online). Available: http://www.tmforum.org/SLA Management/1690/home.html Retrieved: Feb 2010.

MOBILITY

POSTER

A STUDY ON THE USAGE OF MOBILE DEVICES IN COLLABORATIVE ENVIRONMENTS VS DESKTOPS
An Approach based on Flow Experience

Steven Abrantes

Institute Polytechnic of Viseu, Viseu, Portugal
steven@di.estv.ipv.pt

Luis Borges Gouveia

Faculty of Science and Technology, University Fernando Pessoa, Porto, Portugal
lmbg@ufp.edu.pt

Abstract: In order to evaluate the use of mobile devices (laptops) and desktops and the potential of mobile devices in collaborative environments vs desktops, it was performed an experiment involving one hundred and twelve students of higher education. This study has the main objective to validate if the students that use laptops or desktops are in the flow experience and which of them are more in the flow experience. This study is based on the flow experience introduced by Csikszentmihalyi (1975). The main purpose of this study is to establish whether the user is in the flow experience when using Google Groups when using laptops or desktops. In the context of this study, information has been gathered through a survey, applying the five dimensions of the flow state. The sample used consisted on one hundred and twelve students. At the end of the study, after analyzing the gathered information, it was possible to conclude that students have experienced the flow state and it had a positive effect on their learning experiences both by students using laptops or desktops, but having the students that used the laptops a more positive effect in the flow experience than the students that used desktops.

1 INTRODUCTION

Technological applications and the way they are used has advanced in such a way that the manipulation of learning objects is no longer limited to a desktop, but extended to the use of mobile devices (PDA, mobile phone, Smartphone, Laptops, and Tablet PC) to provide a greater range of application and obtain the benefits that mobile computing offers in the education sector. This results in the establishment of a new area of activity, related with the use of mobile technologies in learning, named m-learning.

This educational model based on the use of mobile devices, has been developed over the past few years, resulting in several research projects and some commercial products. Current and past promises of more learning outcomes are needed to be evaluated.

2 M-LEARNING AND MOBILE DEVICES

Quin cited by (Corbeil and Valdes-Corbei, 2007) states that m-learning is the interaction of mobile computing (small applications, portable, and wireless communication devices) with e-learning (learning facilitated and supported through information and communication technologies).

We can see an widespread use of mobile devices in our modern world: mobile phones, PDA's, MP3 players, portable gaming devices, Tablet PCs and laptops, which predominate in our everyday lives.

From children to older people, they are increasingly linked with each other, communicating through communication technologies, something that didn't happen a few years ago.

There are a number of mobile devices that can be considered for an m-learning environment (Corbeil and Valdes-Corbei, 2007): iPod , MP3 Players,

PDA, USB drive, E-Book Readers, Smart Phone, Ultra-Mobile PC (UMPC) and Laptop/Tablet PC. These mobile devices have some advantages and disadvantages (Corbeil and Valdes-Corbeil, 2007). One of the biggest advantages of mobile devices, when compared with desktops, is its ubiquity. With mobile devices people can connect to many kinds of information where they want and whenever they want.

3 THE FLOW EXPERIENCE

An aspect related with the interaction of the users with collaborative environments has to do with the flow experience introduced by Csikszentmihalyi (1975). The flow experience means the sensation that people feel when they are completely involved in what they are doing, that is, people like the experience and want repeat it (Csikszentmihalyi, 1982). This means that for students to be involved with collaborative environments, it is necessary that they presence the flow state.

The theory of the flow allows us to measure the interaction of users with computer systems, verifying if these are more or less playfulness (Trevino and Webster, 1992).

The flow experience is used in this paper to characterize the interaction between the human and the new technologies (Trevino and Webster, 1992).

When one is in the presence of the flow experience, this will bring to the users, a sense of pleasure of what he is doing. This satisfaction will encourage the user to repeat the task again (Webster et al., 1993).

Csikszentmihalyi says that a person who is in the presence of the flow state has the following characteristics(Csikszentmihalyi,1975, Csikszentmihalyi, 1990):

- Clear goals and immediate feedback;
- Equilibrium between the level of challenge and personal skill;
- Merging of action and awareness;
- Focused concentration;
- Sense of potential control;
- Loss of self-consciousness;
- Time distortion;
- Autotelic or self-rewarding experience.

For a person to be in the presence of the flow experience it is necessary a balance between the level of challenge and personal skill (Csikszentmihalyi, 1982) (Figure 1).
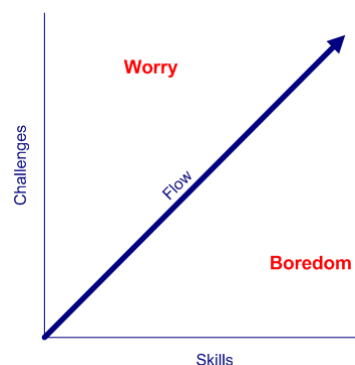
Figure 1: Flow Experience (Csikszentmihalyi, 1982).

The sensation of an excellent experience in the accomplishment of any day by day task is our reason of living. If we do not feel this excellent experience with our everyday tasks, we will question our self, if it is worth living (Csikszentmihalyi, 1982).

Previous researches have used the flow experience to measure playfulness, involvement, satisfaction and other states with the involvement in computational environments (Chen et al., 2000, Ghani and Deshpande, 1994, Novak and Hoffman, 1997, Novak et al., 2000, Trevino and Webster, 1992)

Trevino and Webster (1992) defines four dimensions for the flow experience:

- Control;
- Attention Focus;
- Curiosity;
- Intrinsic Interest.

There is one more dimension, sense of time, that is also important to measure the flow state (McKenna and Lee, 2005) .

People who interact with computers, with an entertainment spirit, transmit a much more positive experience, of those, who are in the computer for obligation (Webster et al., 1993).

4 THE STUDY

To evaluate the flow experience and to verify its occurrence in collaborative tools, an experience was carried through involving one hundred and twelve students from a university school. The main tool used was Google Groups, for this experience. This paper presents the carried through experience, the data obtained, as well as the statistical procedures applied.

After the accomplishment of the project given by the teacher, in which they used Google Groups, the students answered the questions of a survey.

The survey was passed through the Internet with the help of "LimeSurvey". The data collection was performed in the first week of November of 2009.

The Instruments used were Google Groups, Google Docs and Facebook and a survey consisting on some questions, in order to verify, in the end of the study, if the students were in the presence of the flow state. This survey will use the four dimensions: control, attention focus, curiosity and the intrinsic interest (Webster et al., 1993), as well as the dimension sense of time (McKenna and Lee, 2005). Beside these questions, this survey also contains other generic questions. All the related questions from this survey were built on a Likert scale of five points, since one (I totally disagree) up to five (I totally agree). Two questions for each dimension were elaborated.

4.1 Analysis

This study was composed of 78.57% males and 84,82% had ages between sixteen and twenty four years. Most of the students have already used discussion forums in a fairly way.

The majority of the respondents used the laptop (72.32%) to access the tools of the project, followed by the desktop (27,68%).

We verified that Cronbach's alpha is always superior to 0.7, being able to conclude that the data is related to one same dimension, that is, the questions of the survey for the use of Google Groups, allowed us to determine if the individual finds himself in the presence of the flow experience, for students using a laptop or a desktop.

To determine how the variables are correlated with each of the different devices used (laptop and desktop), a correlation matrix was created for both types of the devices, where the correlation coefficient, R, is presented, that is a measure of the linear association between two variables. We can conclude from the correlation analysis that the correlation between the variables, for laptops, has a greater number of variables positively correlated than the desktop.

After the studies mentioned previously, we used the factor analysis in order to reduce the number of variables, both for laptops and desktops.
The extraction of the factors is given by considering the percentage of variance explained by the factors (Table 1 and Table2).

Table 1: Number of factors to be retained (laptop).

| | laptop | | |
| | Initial Eigenvalues | | |
Component	Total	% of Variance	Cumulative %
1	2,371	47,422	47,422
2	,881	17,625	65,047
3	,707	14,136	79,184
4	,631	12,613	91,797
5	,410	8,203	100,000

Table 2: Number of factors to be retained (desktop).

| | desktop | | |
| | Initial Eigenvalues | | |
Component	Total	% of Variance	Cumulative %
1	2,374	47,475	47,475
2	1,053	21,053	68,528
3	,704	14,077	82,604
4	,565	11,301	93,905
5	,305	6,095	100,000

To set the number of components to be retained, we choose, by default, those that have eigenvalues greater than one. If the total variance explained by the factors retained is less than 60%, then, at least, one more factor should always be selected. Thus, for this case study, two factors were retained in each type of device. For the laptop, it appears that the first factor explains 47.422% of the total variation and the second 17.625%, both explaining 65.047% of the total variation that exists in the five original variables. For the desktop, the first factor explains 47.475% and the second 21.053%, explaining both, 68.528% of the total variation.

The matrix of components after rotation (Varimax method) aims to exaggerate the value of the coefficients that relates each variable to the factors retained, so that each variable can be associated with only one factor. The higher the value of the coefficient that relates one variable to a component, the greater is the relationship between them. From this study we have concluded the following for the case of the laptops: Factor group 1: (Intrinsic Interest, Control and Curiosity); Factor group 2: (Attention Focus and Sense of time)
And for the case of the desktops:
Factor group 1: (Attention Focus, Sense of time, Intrinsic Interest and Curiosity) Factor group 2: (Control).

5 CONCLUSIONS

In order to evaluate the use of mobile devices and desktops and the potential of mobile devices in collaborative environments versus desktops, it was performed an experiment involving students of higher education. This study has the main objective to validate if the students that use laptops (mobile device) or desktops are in the flow experience and which of them are more in the flow experience.

The analysis of data allows us to conclude that the majority of the students were males, had ages between sixteen and twenty four years and that most of the students have already used discussion forums.

When going further to the analysis of the data, we verified that the variables described all the same characteristic (threw the determination of the Cronbhach's alpha), that is, the variables describe the flow experience.

We can conclude from the correlation analysis that the correlation between the variables, for laptops, has a greater number of variables positively correlated than the desktop.

From the factor analysis it was possible to isolate two factors that explain the majority of the total variation. Such factors had been Factor group 1: (Intrinsic Interest, Control and Curiosity), Factor group 2: (Attention Focus and Sense of time) for the laptops and Factor group 1: (Attention Focus, Sense of time, Intrinsic Interest and Curiosity) Factor group 2: (Control) for the desktops.

In order to determine the presence of the flow experience for each type of device, it was verified that, on average, the students were above value three (Likert scale of five points), that is, the majority of the students, in each of the different devices (laptop and desktop) used, are in the presence of the flow experience, for the five variables mentioned for this study (attention focus, curiosity, control, intrinsic interest and sense of time). We can also see, that the average of the five variables associated with the flow experience, for students who used the laptops, were greater than those using the desktop to access the tools of the project development.

From this study we can conclude that the flow experience exists for people that use Google Groups, both for people that used the laptop or even the desktop, but having a more positively effect for users of the laptop. Considering that people use mobile device for m-learning and desktops for e-learning, we can conclude that people that use m-learning have a more positive effect on learning than the people that use e-learning.

REFERENCES

Chen, H., Wigand, R. T. & Nilan, M. (2000) Exploring Web users' optimal flow experiences. *Information Technology & People,* 12.

Corbeil, J. R. & Valdes-Corbeil, M. E. (2007) Are You Ready for Mobile Learning? *Educause Quarterly,* 30**,** 51-58.

Csikszentmihalyi, M. (1975) *Beyond Boredom and anxiety,* San Francisco, CA.

Csikszentmihalyi, M. (1982) Towards a Psychology of Optimal Experience. *Annual Review of Personality and Social Psychology.*

Csikszentmihalyi, M. (1990) *The psychology of optimal experience*, Harper Collins.

Ghani, J. & Deshpande, S. (1994) Task Characteristics and the Experience of Optimal Flow in Human-Computer Interaction. *The Journal of Psychology,* 128**,** 381-391.

Mckenna, K. & Lee, S. (2005) A Love Affair with MUDs: Flow and Social Interaction in Multi-UserDungeons.

Novak, T. P. & Hoffman, D. L. (1997) Measuring the Flow Experience Among Web Users. Vanderbilt University.

Novak, T. P., Hoffman, D. L. & Yung, Y. (2000) Measuring the Customer Experience in Online Environments: A Structural Modeling Approach. *Marketing Science,* 19**,** 22-42.

Pereira, P. A. (2002) *Complementos de Estatística.*

Pestana, M. & Gagueiro, J. (2005) *Análise de dados para Ciências Sociais - A complementaridade do SPSS.*

Trevino, L. K. & Webster, J. (1992) Flow in computer-mediated communication. *Communication Research,* 19**,** 539-573.

Webster, J., Trevino, L. K. & Ryan, L. (1993) The dimensionality and correlates of flow in human-computer interaction. *computer game research,* 9**,** 411-426.

BUSINESS COLLABORATION & e-SERVICES

FULL PAPER

GETTING HELP IN A CROWD
A Social Emergency Alert Service

Andreas Geyer-Schulz, Michael Ovelgönne and Andreas C. Sonnenbichler

Information Services and Electronic Markets, Institute of Information Systems and Management
Karlsruhe Institute of Technology (KIT), D-76128 Karlsruhe, Germany
{andreas.geyer-schulz, michael.ovelgoenne, andreas.sonnenbichler}@kit.edu

Abstract: This paper presents the conceptual design of an emergency alert service which addresses the problem of getting help in a crowd. The design is based on the results of more than 30 years of research in social psychology on the bystander effect and it addresses the obstacles for helping by directing an alert directly to the members of the victim's social group who happen to be near the location of the emergency event. The proposed emergency recommender design combines the general availability of geo-location services with social clusters available either from the analysis of social web-sites or from communication analysis. For this purpose, two recently developed innovative clustering methods are used. The feasibility of the design is evaluated by simulating emergency events on the MIT reality mining data set.

1   INTRODUCTION

On Saturday, the 12th of September 2009, 50-year old Dominik Brunner was brutally murdered in a Munich S-Bahn station (Spiegel Online, 2009b). The attack on Dominik Brunner was observed by 15 passengers (Spiegel Online, 2009c) and transmitted and recorded by his mobile-phone on the open police emergency channel (Spiegel Online, 2009a).

This tragic incident reminds social psychologists of the murder of Kitty Genovese on March 13th, 3:20, 1964 in Queens, New York. Kitty Genovese was stalked, stabbed, and sexually assaulted near her own apartment building. During the attack on her she screamed and broke free twice. 38 of her neighbors witnessed the attack, but no one intervened. After 45 minutes one man called the police, but at this point in time Kitty Genovese was dead.

This incident motivated the study of social processes in emergency situations by Darley and Latané (Latané and Darley, 1970) and it points to the shortcomings and problems of real emergency response organizations and their management which very recently have become the object of scientific research e.g. (Yuan and Detlor, 2005), (Faraj and Xiao, 2006), (Chen et al., 2008), and (Comfort et al., 2009). The mobile phone of the victim transmitting and recording to the end confirms Palen and Liu's thesis of the increasing availability of ICT and its use in an emergency by citizens and also their observation, that the traditional hierarchical command-and-control reporting system of emergency response organizations may not be adequate and "does not include built-in considerations for the important roles that members of the public play as participants" (Palen and Liu, 2007, p. 729). Public participation in emergencies and disasters is active and altruistic. First responders are often not the trained professionals of an emergency response organization who are sent to the incident, but the people from the local and surrounding communities. They provide first-aid, transport victims to the hospital, and begin search and rescue (Palen and Liu, 2007, pp. 728-729). In the case of Dominik Brunner we may speculate whether a fast activation of his social community parallel to the emergency call could have saved his life.

The paper starts with a short review of the social processes which Darley and Latené have identified as obstacles for helping in emergencies in section 2. These obstacles have become known as the bystander effect: The more bystanders, the less likely the victim will receive help. We address the bystander effect by a social emergency alert service and discuss how social emergency alert services may help in improving these processes in emergency settings by activating the nearby members of the victim's social network.
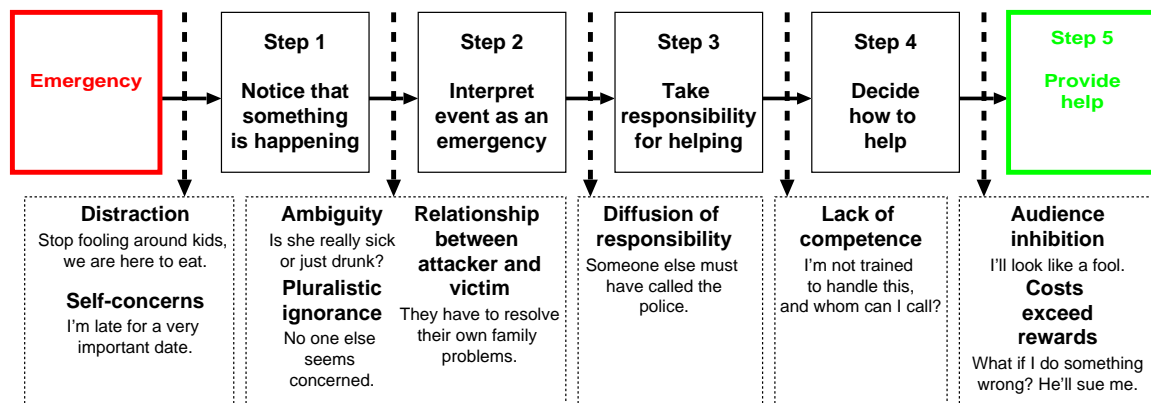
Figure 1: The social help process: 5 steps leading to assistance.

For the process of giving and receiving help, we propose to monitor social interactions and to identify the social clusters of the victims and to locate the nearest members of the social cluster of the victim in an emergency for notification purposes. In section 3 we present details on the social emergency alert service for getting help in a crowd and in section 4 we discuss implementation variants of realizations of such services based on readily available technology by the telecommunication and Internet industry.

The architecture for the prototype of the service which is currently implemented is presented in section 5.

For social cluster identification, recently developed fast algorithms for graph clustering are presented in section 6.

Section 7 of this paper aims at assessing the chances that the social emergency alert service presented has in reality. For this purpose, a first attempt is made to answer three questions which play a crucial role for the success of a social emergency alert service:

1. Has the victim in an emergency a chance to transmit a request for help?

2. Is someone of his social network nearby?

3. Will this person really help?

2 THE UNHELPFUL CROWD: FIVE STEPS TO HELPING IN AN EMERGENCY

The murder of Kitty Genovese in 1964 in Queens, New York, in front of 38 witnesses who did not interfere led Darley and Latané to start investigating the social psychological processes at work in this incident. Their research revealed that the more bystanders, the less likely the victim will be helped. This

is the *bystander effect*: The presence of others inhibits helping. One of the reasons for this is the diffusion of responsibility (Darley and Latané, 1968). Darley and Latané (Latané and Darley, 1970) provided a careful analysis of the process of emergency intervention shown in figure 1. Unfortunately, many obstacles to helping must be overcome. At each step, however, psychological factors are at work which explain why people fail to help:

1. **Noticing.** The presence of others distracts attention from the victim. People who live in big cities may filter out people lying on sidewalks or screams (stimulus overload (Milgram, 1970)).

2. **Interpreting.** People must interpret the meaning of what they perceive. Their perception may be ambiguous: Cries of pain may be taken for laughter, hypoglycemia may be taken for drunkenness (e.g. (Clark and Word, 1972), (Piliavin et al., 1975)). A perceived relationship between attacker and victim may lead the observers to think that everything is OK (Shotland and Straw, 1976). If an emergency happens, the most powerful information available is often what other people do. However, if everybody looks on everybody else to get clues on what to do, the whole group is suffering from pluralistic ignorance and everybody concludes that help is not needed (e.g. (Miller and McFarland, 1987), (Monin and Norton, 2003)).

3. **Take Responsibility.** When help is needed, who is responsible for providing it? The diffusion of responsibility means that people belief that others will or should help. The effect usually is strengthened by anonymity and considerably reduced by a reduction in psychological distance. Groups with members who know each other are more helpful than strangers. See e.g. (Garcia et al., 2002), (Rutkowski et al., 1983), (Baumeister et al., 1995).

4. **Decide how to Help.** Bystanders are more likely to offer direct help if they feel competent to perform the actions required (e.g. (Shotland and Heinold, 1985), (Cramer et al., 1988)).

5. **Provide Help.** Some people may feel too embarrassed to provide help in a public setting (audience inhibition). However, when people think they will be scorned by others for failing to provide help, the presence of an audience will increase their helpful actions. See (Schwartz and Gottlieb, 1980).

In addition, a series of other variables have a high influence on helping behavior as experiments in social psychology have shown: Time pressure reduces the tendency to help (e.g. (Darley and Batson, 1973), (Batson et al., 1978)). Group membership and empathy and attractiveness interact: empathy is a positive predictor for help for in-group members, whereas attractiveness works for out-group members (Stürmer et al., 2005). Group membership positively influences help for in-group members, and the group boundaries can be shifted by proper priming (Levine et al., 2005). Group status and group identification influence the willingness of receiving help (Nadler and Halabi, 2006).

But what can you do to receive help in a crowd? Try to counteract the ambiguity of the situation by making it clear that you need help, and try to reduce the diffusion of responsibility by addressing a specific individual for help, keep eye contact, point or direct requests (e.g. (Moriarty, 1975), (Shotland and Stebbins, 1980)). Consistent with this is a recent study of P. Markey (Markey, 2000) of people in Internet chat rooms: If the number of individuals is large in a chat room, individuals react slower to a plea for help. However, addressing a specific individual by his name leads to considerably faster help and eliminates this effect.

Research on the bystander effect in social psychology showed that even weak social links matter and increase the chance of a victim to receive help considerably. This fact is the main motivation to send alerts to the geographically close members of the victim's social group.

The asymmetric perception of social links (e.g. (Hoser and Geyer-Schulz, 2005)), the role of weak ties, and the cultural norms of the community play a major role in the formation of the social group of the victim. The asymmetric perception of social links implies that a person may not be really aware of possible helpers in his loose social contacts. Taken together with cultural norms, even professional acquaintances are potential helpers. The role of weak ties for networks has been studied by M. Granovetter ((Gra-

novetter, 1973) and (Granovetter, 1982)). In the context of information diffusion on open jobs Granovetter observed that "it is remarkable that people receive crucial information from individuals whose very existence they have forgotten" (Granovetter, 1973, p. 14). This is an indication that an explicit list of emergency contacts provided by the subscriber of such an emergency service will considerably limit the effectiveness of such a service, because of these social phenomena.

3 GETTING HELP IN A CROWD - A SOCIAL EMERGENCY ALERT SERVICE

Getting help in a crowd is a social emergency alert service selecting nearby members of the social group of the victim and notifying them about the victim's need for help and the victim's location. With this service we aim to activate the locally available social network of the victim and to eliminate the bystander effect.

The UML sequence diagram depicted in figure 2 shows the generic process in an emergency incident. It is designed on a high-level, abstract way allowing a variety of industrial implementations. We will address this issue in section 4.

An emergency notification is submitted by the victim by starting an application on his mobile device (`LaunchEmergenyApplication` in figure 2) e.g. by pressing the help-button shown in figure 5. The application retrieves the current geo-position. Both, the ID of the emergency caller and his geo-location are then transmitted to the emergency alert service (`EmergencyAlert` in figure figure 2).

For discovering the most likely person to help in the victim's social network, his social network has to be known and possible helpers identified (`GetHelperCluster` in figure 2). The network is either built from social interaction data from e-mail, sms, phone, and mobiles where the number of interactions is taken as an indication of social nearness or from social web sites as for example Facebook or Xing. However, the number of social interactions may be ambiguous as a recent incident (Rungg, 2010) of a woman threatened by her former husband with Googles Buzz has shown: So the possibility to check for such unwanted relations must be provided for the participants of the emergency alert service (`FilterCandidates` in figure 2). Usually social networks tend to be very large. As emergencies are often time-critical, it might take too long to calculate such a network on-the-fly. The network is pre-built and up-
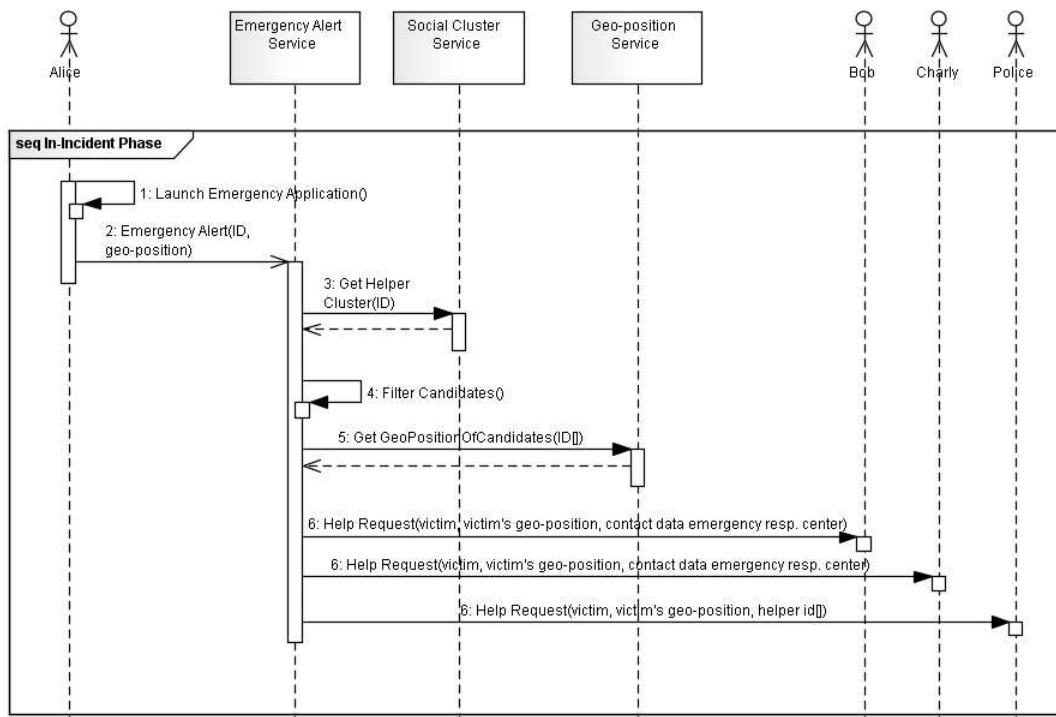
Figure 2: Sequence Diagram of a Generic Social Emergency Alert Service (in-incident).

dated regularly for all service subscribers. To find out which persons in the social network are likely to help, a clustering of the network is performed. Details for this clustering are addressed in section 6.

Next, the current geo-position of the candidates is retrieved from a geo-position service (GetGeoPositionofCandidates in figure 2). The alert service uses the geo-data as a filter on the victim's social cluster to find out, who of the possible helpers is locally close enough. Section 7 deals with details of having at least one member of the victim's social cluster in range.

The possible helpers in range and the emergency response center are informed about the emergency situation of the victim (several invocations of HelpRequest in figure 2). All possible helpers in range are informed simultaneously. The victim's name, his geo-location and the shortest route as well as the contact data for the emergency response center are provided.

Finally, possible helpers and the emergency response experts at the police's emergency response center may communicate, because of the information forwarded by the emergency alert service (not shown in figure 2). This facility has the potential of providing expert guidance to the socially close first responders on the scene. However, it also reveals the identity of potential helpers to the emergency response center.

The privacy impact of this must be addressed for such a service.

In section 2 obstacles to the five steps leading to assistance have been described. The emergency alert service presented addresses these obstacles directly:

1. The *distraction* obstacles can be avoided by noticing, that an emergency incident takes place: Clear signal words are part of the personal message to the helpers. This makes it obvious, an emergency case is happening and this is made clear to the helper.

2. *Self concerns* are also addressed by the service: Since the potential helper is directly addressed and others know this from the incident protocol, social norms lead to pressure to help.

3. *Ambiguity* is by-passed by the clear and unmistakable help request sent to the helpers.

4. As this message is personal, *pluralistic ignorance* is eliminated. Experimental evidence for these effects in internet chat rooms is provided by P. Markey (Markey, 2000).

5. The *relationship between attacker and victim* can not lead the helper to overlook the emergency event, because of the unambiguous emergency message.

6. *Diffusion of responsibility* is also reduced, since the emergency alert message is directly and per-

sonally addressed to the helper. Because of this, he is responsible and because of the incident protocol, others will know this and hold him responsible.

7. *Lack of competence* may be addressed by providing fast expert backup for helpers from the police emergency response center.

8. Expert guidance of how to help also addresses the problems of *audience inhibition* and *costs exceed rewards*.

4 IMPLEMENTATION VARIANTS

In this section we present implementation examples how the *Social Emergency Alert Service* can be implemented in an industrial environment.

4.1 Emergency Alerting

To be able to use the service, the user has to possess a mobile device (e.g. a mobile phone). He can then subscribe to the service. In case of an emergency, he starts an application on his mobile device. Of course, the start of the application must be made simple and fast, as we do not expect it likely that the victim is able to deal with complex applications in an emergency situation. For the implementation third-party platforms like Android can be used. Android (Android, 2010) is a mostly free and open-source OS platform developed and driven by the Open Handset Alliance (Open Handset Alliance, 2010). Further platforms like Apples iPhone may be supported as well. Special mobile devices combined with body-sensors, e.g. for elderly people, linking the start of the application to a hardwired button can be offered.

4.2 Geo-position Service

Geo-positions of both the victim and all possible helpers of the victim's social network need to be calculated. Many of todays mobile phones are able to calculate their geo-position by GPS (Global Positioning System). The service Google Latitude is an example for a service that users can publish their current geo-position and share it with friends. If a mobile device does not include such a feature, several alternative techniques have been described and implemented. Even speed vectors can be calculated (for example (Kikiras and Drakoulis, 2004; Borkowski and Lempiainen, 2006)). By this, the expected geo-position of somebody moving in a train can be found out.

4.3 Building Social Networks and Identifying Possible Helpers

For discovering the most likely person to help in an emergency case the social network has to be known. We present three possible realizations.

The social network can be built by monitoring outgoing and incoming calls on the mobile device of the subscriber. The emergency alert application running on the mobile device collects this call data, pools it and regularly (e.g. once a week) transmits it via HTTP to the social cluster service (see figure 2). There the call logs of all service subscribers are combined and the network is calculated: Telephone numbers are represented as nodes, the calls are weighted ties. Each call strengthens a tie. The advantage of this solution is, that the network is independent from the telephone provider. It works depending just on the emergency alert application. The disadvantage is, that the calculated social network consists only of subscribers and their direct connections. Ties between non-subscribers can not be observed technically.

Alternatively not the mobile devices monitor the calls, but connection records from telephone providers are used. Connection records are stored for billing purposes. In the European Union an directive forces the provider to save call logs from six month up to two years (EU, 2006). These connection data can be used to calculate the social networks. Every connection is represented by a tie between the calling parties (more concrete, their telephone numbers) as nodes. Of course, the resulting network will be huge. In section 7 we will show that feasible cluster algorithms for such huge data sets exist. The advantage of this solution is, that much more network data can be collected so that the problem of missing links is smaller. On the other hand this alternative can only be realized if the calling logs are available to the emergency alert provider. As we do not expect network providers to give such information away, the most likely approach for this alternative is, if the network provider is identical with the emergency alert service provider. The provider can then use the service as an additional opt-in feature. Another disadvantage is, that one network provider will probably not exchange network or call log data with other providers. By this the social network is limited to the customers of the provider plus their direct links.

As an additional feature for both alternatives address books in the customers' devices can be used to group telephone numbers. Different telephone numbers of one person can be combined and fused to one node in the social network.

A third approach to build the social network is

to cooperate with existing social network platforms. Data from Facebook, Xing, Myspace can be used. The advantage of this solution is, that no subscriber or network provider boundaries exist. The disadvantage is, that people tend to accept more 'friends' in social platforms than they would accept offline. Additionally most of these platforms do not weight their ties, which makes it difficult to apply cluster algorithms.

In practice all three alternatives used to build a social network as discussed above can be complemented with a list of emergency contacts provided by each subscriber and, if available, with a list of dedicated helpers for an event or for a community. In a German small rural community, the community's voluntary fire-fighters are an example of such a community. We expect, that people in the same social group are likely to help each other.

5 ARCHITECTURE OF THE PROTOTYPE

In figure 3 we present the architecture of a prototype of the service. The mobile application (*Emergency Alert Widget*, *GPS-Push Service*, *Communication Capture Service* in figure 3) is implemented on a Motorola Milestone smart phone using Android as OS. The *Service Provider* is realized in Java using Apache Service Mix as OSGi [1] provider.

Figure 4 depicts the sequence diagram of the prototype. Each subscriber is identified by his unique telephone number (ID). In the *Pre-Incidence Phase* each subscriber submits his geo-location data to the *GPS Cache* (step 1). We use a *GPS Push Service* running as background service on the Android smart phone to transmit the GPS coordinates via REST to the *GPS Cache*. The transmission is done on a regular basis, e.g. every 10 minutes and if the position changed by more than 10 meters. The *GPS Cache* stores the latest transmitted geo-position of each subscriber in a database. The *Communication Capture Service* runs on the mobile device. It monitors all ingoing and outgoing calls and messages. This communication data is collected in a local cache. Once a day the smart phone submits the cached communication data via SOAP to the *Communication Collector* (step 2). The *Social Cluster Calculator* pulls new data from the *Communication Collector* on a daily basis. It builds a communication network, clusters the groups and stores the results in the *Social Group Provider's*

---

[1] The Open Services Gateway initiative is a module system and service platform for Java. For details see http://www.osgi.org

database (step 3 and 4). OSGi is used as the internal Service Provider's protocol.

In the *In-Incident Phase* the user in need (in our example this is Alice) starts a widget on her smart phone (see figure 5). A emergency alert is transmitted via REST to the *Emergency Alert Service* (step A). The service pulls the helpers candidate list from the *Social Group Provider* for Alice via OSGi (step B). For each candidate the *GPS Cache* is searched for the latest geo-position data of the candidate (step C). If close enough, a MMS is sent out to the helper with information about the victim and the victims geo-position (depicted on a map).

Note, that Figure 4 is restricted to a proper emergency alert. False alerts can be revoked by a similar process (not shown in Figure 4) which is password protected. However, pragmatically a set of passwords is provided which act as silent signals. A small solution consists of three passwords, the first signaling a false alarm, the second signaling that the victim is forced to revoke the alarm, and the third that there is danger for the helpers.

6 IDENTIFYING SOCIAL GROUPS BY CLUSTERING SOCIAL NETWORKS

Calling persons willing to help is crucial for the proposed system but their identification is not trivial. Communication networks or 'friend' networks of online social network sites usually contain many links that do not result from close personal relations. Links may connect business partners or co-workers. On social network sites people 'friend' others they rarely know. Therefore, identifying social groups is an approach to separate close personal contacts from other distant contacts that are less willing to help in a case of emergency.

The appropriate cluster algorithms depend on the network that needs to be analyzed. All algorithms need to be highly efficient as the mentioned networks are very huge (several million vertices). From communication data weighted networks could be created where an edge connects caller and the callee respectively sender and receiver of a text message. The edges can be weighted by the number of calls or messages. Walk context clustering is a suitable method for this kind of network. It generates overlapping clusters and can reflect that people might have several groups of close contacts (family, friends, neighbors) that are almost not connected with each other.

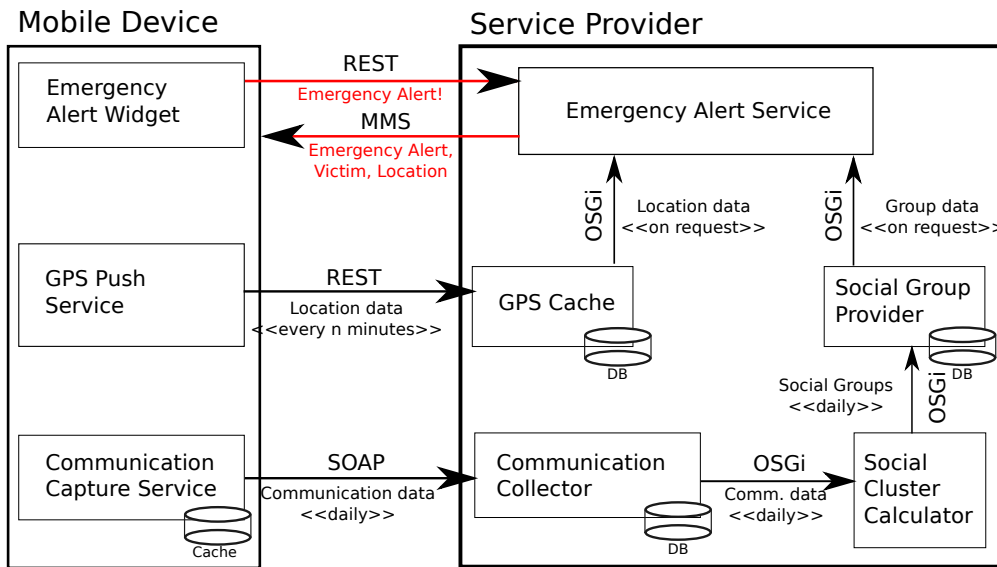Cluster algorithms based on optimizing the mod-

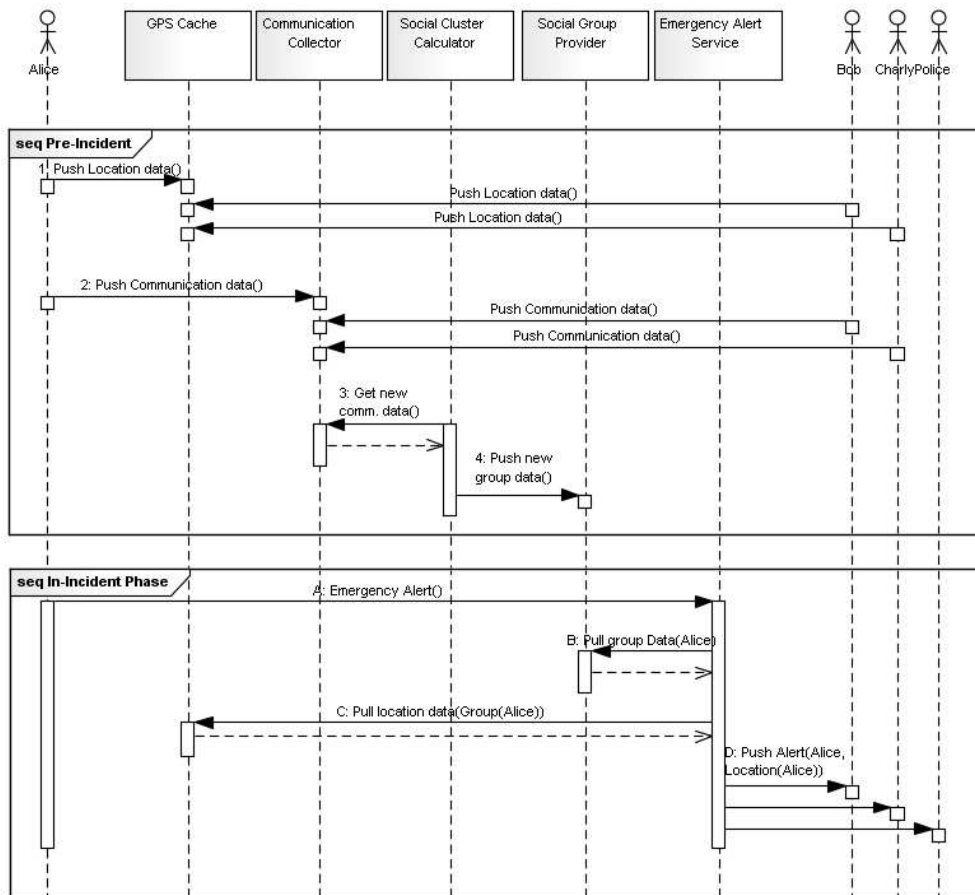Figure 3: Architecture of the Prototype.

Figure 4: Sequence Diagram of the Prototype.

ularity measure can not reflect the multitude of social groups one person is integrated in. However, they are able to deal with unweighted networks like the friendship networks of social network sites.
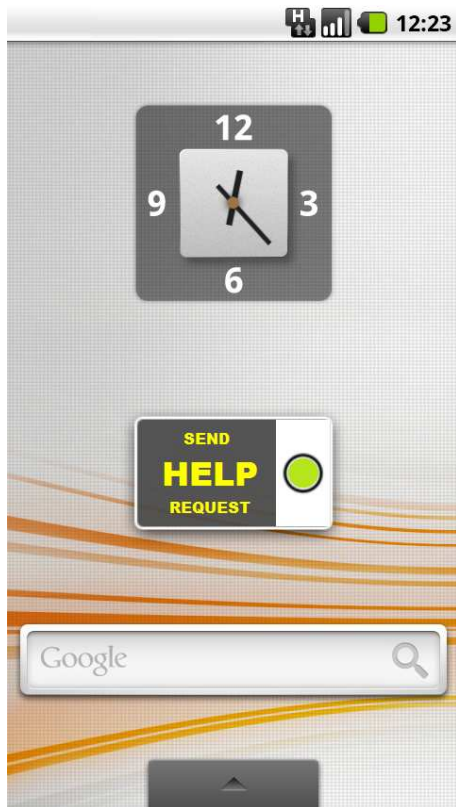
Figure 5: Alert Widget on Phone.

6.1 Walk Context Clustering

Walk context clustering consists of two stages (see algorithm 1). In the walk stage, a set of restricted random walks is generated by starting a number of walks at each vertex and repeatedly choosing the following vertices randomly among those vertices that are linked by an edge which has a higher weight than the previously taken one (see figure 6).

In the cluster construction stage, clusters get generated from the walks. Walk context clustering assigns a vertex to the cluster of another vertex if both are part of the same walk. A level parameter l specifies the fraction of vertices at the beginning of a walk that are disregarded. The later a pair of vertices appears in a walk the stronger is their connection. The interesting feature of walk-context clustering is that the closeness of two persons can be measured by the maximal level that assigns one person to the cluster of the other one.

A recently developed database-backed update algorithm for the walk stage maintains asymptotically optimal clusters in near real-time (< 0.2 sec for a single update on graphs with approximately 500000 nodes and 20 million edges) (Franke and Geyer-Schulz, 2009).

Algorithm 1: Walk context clustering.

Data: undirected, weighted graph $G = (V, E)$, constant p

▼ **Walk generation**
  $walkSet \leftarrow \emptyset$ **forall** $v \in V$ **do**
    for $counter \leftarrow 1$ to p **do**
      $walk \leftarrow (i)$; $last \leftarrow 0$; $i \leftarrow v$ **while** $N = \{x | \omega_{ix} > last\} \neq \emptyset$ **do**
        $j \leftarrow$ random element of N $last \leftarrow \omega_{ij}$ append j to walk w $i \leftarrow j$
    $walkSet \leftarrow walklist \cup walk$

Data: walkSet ws, vertex v, level l
▼ **Cluster Construction**
  $cluster \leftarrow \emptyset$ **forall** $w \in ws | v \in w \wedge pos(v,w) > l$ **do**
    forall $x \in w | pos(x,w) > l$ **do**
      $cluster \leftarrow cluster \cup x$
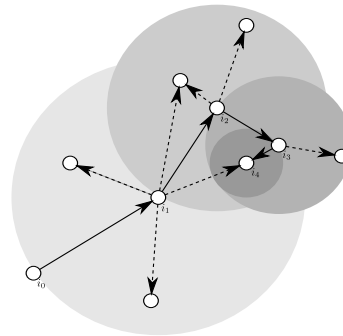
Figure 6: Narrowing search space for successive vertices of the restricted random walk algorithm. The walks terminates when no neighbor is within the search space. Solid arrows symbolize used transitions. Dashed arrows symbolize links to possible successors that have not been chosen by the random process.

6.2 Modularity Clustering

Newman and Girvan (Newman and Girvan, 2004) proposed a quality function for graph clusterings. The modularity Q of a clustering $C = C_1, ..., C_z$ ($\forall i, j : C_i \cap C_j = \emptyset$ and $\cup_i C_i = V$) of a graph $G = (V, E)$ is

$$Q = \sum_i (e_{ii} - a_i^2) \qquad (1)$$

with $e_{ij} = \frac{\{\{v_a, v_b\} \in E | v_a \in C_i \wedge v_b \in C_j\}}{2|E|}$ and $a_i = \sum_j e_{ij}$.

The randomized greedy modularity clustering algorithm (Ovelgönne and Geyer-Schulz, 2009) (see algorithm 2) is a fast agglomerative hierarchical clustering algorithm. It places each vertex in a separate cluster and builds the complete dendrogram by repeatedly merging clusters. In each step it randomly selects k clusters and searches among the clusters and

their neighbors for the pair that yields in the highest increase in modularity. The ΔQ of the merge of two clusters i and j is $\Delta Q(i,j) = e_{ij} + e_{ji} - 2a_i a_j = 2(e_{ij} - a_i a_j)$. This algorithm is able to process even very large networks in reasonable time. A graph with about 300.000 vertices and 1 million edges is clustered in roughly 10 seconds.

Algorithm 2: Randomized greedy modularity clustering.

Data: undirected, connected graph $G = (V, E)$, constant k
▼ **Initialize**
  forall $v \in V$ **do**
    forall *neighbours n of v* **do**
      $e[v,n] \leftarrow 1/(2 * edgecount)$
    $a[v] \leftarrow rowsum(e[v])$

▼ **Build Dendrogram (Randomized Greedy)**
  for *i = 1 to rank(e)-1* **do**
    maxDeltaQ $\leftarrow -\infty$ **for** *j = 1 to k* **do** search among k communities for best join
      $c1 \leftarrow$ random community **for** *all communities c2 connected to c1* **do**
        $deltaQ \leftarrow 2(e[c1,c2] - (a[c1] * a[c2]))$ **if** $deltaQ > maxDeltaQ$ **then**
          $maxDeltaQ \leftarrow deltaQ$ $nextjoin \leftarrow (c1, c2)$

    join(nextjoin)

---

7 ASSESSMENT OF BENEFITS

7.1 Ability to Transmit Request for Help

Emergency situations can result of various incidents, e.g. crime, accidents, medical emergencies. By their nature, accidents happen unexpected and sudden. Some medical emergencies as heart attacks do as well. The ability to make an emergency call in these cases will depend on the physical condition of the person in need.

For victims of violent crime their ability to send an emergency signal depends on the progress of crime. In 25%, respectively 22%, of the robberies analyzed by Smith (Smith, 2003) instant violence or attempts to snatch property don't give the opportunity to call for help. However, in 37% of the robberies the offender approached the victim and demanded money or valuables without immediate violence. In many cases later victims were also aware of an upcoming threat. In those cases it would be possible to send an emergency message.

7.2 Chance for Nearby Help

The helpfulness of the proposed system depends on the availability of close social contacts in the proximity of the site of the emergency. The actual number of persons in ones proximity in general and the number of close contacts with a particular motivation to help will surely depend on several factors, e.g. place and time.

To gain insight into the availability of potential help the MIT reality mining experiment (Eagle and Pentland, 2005) provides an interesting data set. For this experiment a group of 100 persons (75 students or faculty of the MIT Media Laboratory and 25 students of the MIT Sloan Business School) has been equipped with smart phones with special software applications preinstalled. These applications recorded phone numbers of incoming and outgoing calls, text messages, and the id of the cellular tower the phones were connected to during one academic year.

The phone call and text message data from the reality mining data set can be used to construct a communication network. Based on the assumption that the communication intensity of two people reflects the closeness of their relation, social groups can be identified by clustering this weighted network. The availability of nearby help from within the social group of a person in need can be estimated by the number of persons from the same social group whose phones are connected to the same cellular tower as the person in need.

For three consecutive months all communication prior to the specific month has been used to build an undirected, weighted communication network. The edge weights are the number of communication events (calls, text messages). This network has been clustered by the walk context cluster algorithm with the level parameter set to 0.8. The generated clusters had an average size of about 5.

The MIT reality mining data set contains a history of time-stamped connection records of the participating persons. For each month 1000 connection records have been randomly selected. Then, for each persons in a connection record the number of persons of his social cluster who have been connected to the same cellular tower at the same time have been counted. This simulation showed that on average more than one close fellow student was available for help at any time (see table 1). E.g. for September 2004 the probability of having at least one person from one's social group in the proximity was 78% - independent of time of the day.

This is just a basic assessment for a particular group of people that has several shortcomings. Due to the lack of more detailed data it was not possible to as-

Table 1: Average number of other persons of same social group connected to same cellular tower for 1000 randomly selected persons and points of time.

	Day (6am-6pm)		Evening (6pm-11pm)		Night (11pm-6am)	
	Group	Others	Group	Others	Group	Others
09/2004	1.16	3.74	1.17	2.73	0.95	1.51
10/2004	1.33	3.12	1.29	3.4	1.14	1.82
11/2004	2.2	3.18	1.53	2.16	1.76	1.03

sess if other social contacts than fellow students were available for help. Proximity could just be estimated by radio cells which have a radius of a few hundred meters in urban areas and a radius up to about 35 km in rural areas. But it is fair to regard the results as an indication that people living in urban areas will usually have at least one of their social contacts in their proximity.

7.3 Chance to Actually Receive Help

Once a request for help has been transmitted to a potential helper in the proximity of the site of the emergency the chance to actually receive help depends on the willingness and the ability of the informed persons to get to that place. The ability to get to a specific site can be supported in various ways. For example a map and route directions could be displayed together with the emergency message. The research of Markey (Markey, 2000) showed that also in computer-mediated communication settings help requests that were directed to specific persons had a high probability to receive fast response and the bystander effect was virtually eliminated.

8 CONCLUSIONS

In this paper a novel emergency alert service has been introduced which addresses all obstacles to providing help identified in the social emergency intervention process discussed in section 2. The service is designed to reduce psychological barriers that result in a bystander effect and inhibit effective help for persons in need. The analysis of emergency situations and whereabouts of persons in relation to their respective social group suggest that the described service can actually provide a benefit in practice. A first prototype of this service is currently under development.

The main emphasis of this paper is on the reduction of the bystander effect. However, the following challenges which are beyond the scope of this paper are examples of what must be addressed thoroughly for concrete industrial service offerings:

1. Legal issues: The service raises e.g. the problem

that potential helpers become liable to help and failure to do so may be prosecuted.

2. Privacy: The service should be designed in order to minimize the intrusion of privacy of service-subscribers.

3. Emergency dialog: The emergency dialog could be further automated e.g. by providing an automatic classification of the incident type.

4. Geo-positioning problems: Geo-positioning is still problematic in large buildings, tunnels, subterranean areas (e.g. subway). Enhancements could be based e.g. by embedding geo-position senders in such structures or by image recognition techniques which exploit public geo-coded images of such spaces.

ACKNOWLEDGEMENTS

REFERENCES

Android (2010). Android. http://www.android.com/.

Batson, C. D., Cochran, P. J., Biederman, M. F., Blosser, J. L., Ryan, M. J., and Vogt, B. (1978). Failure to help when in a hurry: Callousness or conflict? *Personality and Social Psychology Bulletin*, 4:97 – 101.

Baumeister, R. F., Chesner, S. P., Sanders, P. S., and Tice, D. M. (1995). Who's in charge here? group leaders do lend help in emergencies. *Personality and Social Psychology Bulletin*, 14:17 – 22.

Borkowski, J. and Lempiainen, J. (2006). Practical network-based techniques for mobile positioning in UMTS. *EURASIP J. Appl. Signal Process.*, 2006:149–149.

Chen, R., Sharman, R., Rao, H. R., and Upadhyaya, S. J. (2008). Coordination in emergency response management. *Communications of the ACM*, 51(5):66 – 73.

Clark, R. and Word, L. E. (1972). Why don't bystanders help? because of ambiguity? *Journal of Personality and Social Psychology*, 24:392 – 400.

Comfort, L. K., Kilkon, K., and Zagorecki, A. (2009). Co-ordination in rapidly evolving disaster response systems: The role of information. *American Behavioral Scientist*, 48(3):295 – 313.

Cramer, R. E., McMaster, M. R., Bartell, P., and Dragna, M. (1988). Subject competence and the minimization of the bystander effect. *Journal of Applied Social Psychology*, 18:1133 – 1148.

Darley, J. M. and Batson, C. (1973). "From Jerusalem to Jericho": A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 27(1):100 – 108.

Darley, J. M. and Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8(4):377 – 383.

Eagle, N. and Pentland, A. (2005). Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, 9:1 – 14.

EU (2006). Directive 2006/24/EC of the European Parliament and of the Council on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC.

Faraj, S. and Xiao, Y. (2006). Coordination in fast-response organizations. *Management Science*, 52(8):1155 – 1169.

Franke, M. and Geyer-Schulz, A. (2009). An update algorithm for restricted random walk clustering for dynamic data sets. *Advances in Data Analysis and Classification*, 3(1):63 – 92.

Garcia, S. M., Weaver, K., Moskowitz, G. B., and Darley, J. M. (2002). Crowded minds: The implicit bystander effect. *Journal of Personality and Social Psychology*, 83:843 – 853.

Granovetter, M. (1982). The strength of weak ties: a network theory revisited. chapter 5, pages 105 – 130. Sage Publications, Inc., Beverly Hills, California.

Granovetter, M. S. (1973). The strength of weak ties. *The American Journal of Sociology*, 78(6):1360 – 1380.

Hoser, B. and Geyer-Schulz, A. (2005). Eigenspectral analysis of hermitian adjacency matrices for the analysis of group substructures. *The Journal of Mathematical Sociology*, 29(4):265 – 294.

Kikiras, P. and Drakoulis, D. (2004). An integrated approach for the estimation of mobile subscriber geolocation. *Wireless Personal Communications*, 30(2):217–231.

Latané, B. and Darley, J. (1970). *The Unresponsive Bystander: Why doesn't he help?* Appleton-Century-Crofts, New York.

Levine, M., Prosser, A., Evans, D., and Reicher, S. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior. *Personality and Social Psychology Bulletin*, 31(4):443 – 453.

Markey, P. M. (2000). Bystander intervention in computer-mediated communication. *Computers in Human Behavior*, 16:183 – 188.

Milgram, S. (1970). The experience of living in cities. *Science*, 167:1461 – 1468.

Miller, D. T. and McFarland, C. (1987). Pluralistic ignorance: When similarity is interpreted as dissimiliarity. *Journal of Personality and Social Psychology*, 53:298 – 305.

Monin, B. and Norton, M. (2003). Perceptions of a fluid consensus: Uniqueness bias, false consensus, false polarization, and pluralistic ignorance in a water conservation crisis. *Personality and Social Psychology Bulletin*, 29:559 – 567.

Moriarty, T. (1975). Crime, commitment, and the responsive bystander: Two field experiments. *Journal of Personality and Social Psychology*, 31:370 – 376.

Nadler, A. and Halabi, S. (2006). Intergroup helping as status relations: Effects of status stability, identification, and type of help on receptivity to high-status group's help. *Journal of Personality and Social Psychology*, 91(1):97 – 110.

Newman, M. E. J. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2):026113.

Open Handset Alliance (2010). Open handset alliance. http://www.openhandsetalliance.com/.

Ovelgönne, M. and Geyer-Schulz, A. (2009). Randomized greedy modularity clustering. Technical report, Karlsruhe.

Palen, L. and Liu, S. B. (2007). Citizen communication in crisis: Anticipating a future of ict-supported public participation. In *CHI 2007 Proceedings*, pages 727 – 736. ACM.

Piliavin, I. M., Piliavin, J. A., and Rodin, J. (1975). Costs, diffusion, and the stigmatized victim. *Journal of Personality and Social Psychology*, 32(3):429 – 438.

Rungg, A. (2010). Privatspäre 2.0. Technical report, Financial Times Deutschland, 17.2.2010, p.25.

Rutkowski, G. K., Gruder, C. L., and Romer, D. (1983). Group cohesiveness, social norms, and bystander intervention. *Journal of Personality and Social Psychology*, 44:545 – 552.

Schwartz, S. H. and Gottlieb, A. (1980). Bystander anonymity and reaction to emergencies. *Journal of Personality and Social Psychology*, 39:418 – 430.

Shotland, R. L. and Heinold, W. D. (1985). Bystander response to arterial bleeding: Helping skills, the decision-making process, and differentiating the helping response. *Journal of Personality and Social Psychology*, 49:347 – 356.

Shotland, R. L. and Stebbins, C. A. (1980). Bystander response to rape: Can a victim attract help? *Journal of Applied Social Psychology*, 10:510 – 527.

Shotland, R. L. and Straw, M. K. (1976). Bystander response to an assault: When a man attacks a woman. *Journal of Personality and Social Psychology*, 34:990 – 999.

Smith, J. (2003). The nature of personal robbery. Home Office Research Study 254, Home Office Research, Development and Statictics Directorate.

Spiegel Online (2009a). Notruf-Mitschnitt belegt Brutalität der S-Bahn-Prügler. Technical report, Spiegel Online.

Spiegel Online (2009b). S-Bahn-Fahrgast stirbt nach brutaler Attacke. Technical report, Spiegel Online.

Spiegel Online (2009c). Tödlicher Angriff: Mehrere Menschen beobachteten Münchner S-Bahn-Attacke. Technical report, Spiegel Online.

Stürmer, S., Snyder, M., and Omoto, A. M. (2005). Prosocial emotions and helping: The moderating role of group membership. *Journal of Personality and Social Psychology*, 88(3):532 – 546.

Yuan, Y. and Detlor, B. (2005). Intelligent mobile crisis response systems. *Communications of the ACM*, 48(2):95 – 98.

SHORT PAPER

SEMANTIC INTEROPERABILITY OF E-SERVICES
IN COLLABORATIVE NETWORKED ORGANIZATIONS

Nikos Papageorgiou, Yiannis Verginadis

Institute of Communications and Computer Systems, National Technical University of Athens
9 Iroon Polytechniou Str., Athens, Greece
{npapag, jverg}@mail.ntua.gr

Dimitris Apostolou

Informatics Department, University of Piraeus, Karaoli & Dimitriou Str., Piraeus, Athens, Greece
dapost@unipi.gr

Gregoris Mentzas

Institute of Communications and Computer Systems, National Technical University of Athens
9 Iroon Polytechniou Str., Athens, Greece
gmentzas@mail.ntua.gr

Abstract: Semantic interoperability is a crucial issue in enterprises when they participate in Virtual Organizations (VOs). Addressing semantic heterogeneities, detected in VOs, aims to ensure that the meaning of information exchanged is interpreted in the same way by all communicating parties and their systems. In this paper we examine how ontologies can be employed by a system of e-services for delivering interoperability to enterprises, independent of particular IT deployments. In order to support interoperability service utilities in VOs, this paper presents a top-level ontology for collaborative networked organizations (code named OCEAN). The OCEAN ontology is designed as a lightweight top-level ontology that provides a common terminological reference for e-services supporting VO collaborations. We demonstrate how the usage of OCEAN enables e-service interoperability in knowledge-intensive collaborations presenting concrete examples from the pharmaceutical industry.

1 INTRODUCTION

Enterprises wishing to take part in collaborative networks participate in formations often referred to as Virtual Organizations (VOs) (Davidow et al., 1992), (Mowshowitz, 1997). A VO is a short-term association with a specific goal of acquiring and fulfilling a collaboration opportunity. A key underpinning of VOs is the logical separation of VO members' requirements (e.g., requests for information, advice, or transactions) from satisfiers (e.g., information services, collaboration services, or transactional services) (Mowshowitz, 1997). Having such a capability allows management to continually examine service requirements, scan for matching service offerings and switch the assignment of satisfiers to requirements so as to optimize performance on the basis of explicit criteria such as

reducing service delivery costs or improving service quality. Since each VO member undertakes particular sub-processes in the joint effort, information and services for enabling knowledge-based collaboration should be available in an interoperable way. Towards this end, adequate semantic interoperability has to be established by means of a common frame of reference or at least a common terminology (Chituc et al., 2008).

Advances in Semantic Web (Berneers-Lee, 2007) technologies, which enable machines to process and reason about resources in support of businesses interactions, have paved the way for ontology-based platforms enabling semantic interoperability between heterogeneous information systems. In this paper we examine how ontologies can be employed by a system of e-services for delivering interoperability to enterprises,

independent of particular IT deployments.

The main contributions of the paper are the following. First, the paper proposes an ontology representing VO objects, processes, roles and relationships as a formal framework for enabling resolution of semantic heterogeneities. Second, the paper presents the methodology that we have used for customizing the ontology for the particular needs of VOs and for achieving consensus of the shared conceptualization of a VO among participants. Third, concrete examples from the pharmaceutical industry are used to demonstrate the applicability and benefits of the proposed ontology and architecture in supporting VO collaboration.

The remaining of the paper is organized as follows: In section 2 we discuss the emerging need for supporting semantic interoperability in VOs, while in section 3 the methodology used for the development of the OCEAN ontology are presented. In section 4 we present the main concepts of this ontology. Finally, our paper concludes with the application of our work in the pharmaceutical industry in section 5 and with a discussion on the research implications and conclusions in section 6.

2 SEMANTIC INTEROPERABILITY IN VIRTUAL ORGANIZATIONS

2.1 Interoperability and Ontologies

In the context of networked enterprises (i.e. enterprises that participate in a VO), interoperability refers to the ability of interactions (exchange of information and e-services) between enterprise systems. The Enterprise Interoperability Research Roadmap – EIRR argues that interoperability of enterprises in future business ecosystems will be a utility-like capability that enterprises can invoke on the fly in support of their business activities. The European Commission uses the term Interoperability Service Utility (ISU) to denote a basic "infrastructure" that supports information exchange between diverse knowledge sources, software applications, and Web Services.

Current interoperability solutions are often oriented toward integration of data required for executing a common business goal, often specified in terms of a contract. Protocols and standards such as ebXML(2009), Electronic Data Interchange (EDI, 2009), and RosetaNet (2009) have been enablers for the progress made in the ability to integrate heterogeneous information and data.

But, semantic interoperability aims to achieve a more ambitious goal, that is to assure that the meaning of the information exchanged (e.g., business documents, messages) is interpreted in the same way by the communicating systems (Chituc et al., 2008). For addressing semantic heterogeneity it is essential that the semantic definitions of the knowledge objects, processes, roles and relationships within VOs are defined based on a mathematically rigorous ontological foundation (Lin et al., 2007). Moreover, as VO members might come from different fields or have different professional backgrounds, it is necessary to introduce a mechanism to share common understanding of knowledge, and to agree on a controlled vocabulary. An ontology provides a representation of knowledge, which can be used in order to facilitate the comprehension of concepts and relationships in a given domain, the communication between VO members by making the domain assumptions explicit and the resolution of semantic heterogeneities between VO systems.

2.2 Existing Approaches

Among the wide spectrum of approaches which differ in the amount of information and specificity, four categories of approaches can be distinguished for developing ontologies i.e., top level, domain, task and application level ontologies (Huang et al., 2007), (Rajpathak et al., 2006), (Andersson, 2006). Top level ontologies are used to represent the building blocks for a particular domain and basically constitute the first step toward knowledge representation for a domain. Basically, this kind of ontology is limited to concepts that are meta, generic, abstract and philosophical, and therefore are general enough to address (at a high level) a broad range of domain areas. In the last decade, many projects aimed at creating top level ontologies for different purposes: word net (Fellbaum,1998), SUMO (Niles et al., 2001), DOLCE (Gangemi et al., 2002), AIAI Enterprise Ontology (Uschold et al., 1998), PROTON (Kiryakov, 2006), ECOLEAD (Plisson et al., 2007) and the Business Management Ontology (BMO), TOVE ontologies for enterprise modeling (TOVE), and the DIP Business Data Ontology (DIP) and ontologies for enterprise interoperability (Ruokolainen et al., 2007), (Castano et al., 2006).

Among these most relevant to our work is the ECOLEAD ontology which proposes an ontology for Virtual Breeding Environments, which are long-

term associations of enterprises that have the potential and the will to form a VO. The OCEAN ontology builds upon and extends the ECOLEAD ontology to cover the creation, operation and termination phases of VOs. In particular, we focus on knowledge-oriented collaboration within VOs and subsequently OCEAN aims to enable interoperability of systems providing e-services for enabling knowledge-based collaboration.

3   METHODOLOGY

Among the various ontology development methods that have been proposed (Cristani et al., 2005), we opted for a collaborative method because it addresses the objective of achieving a shared representation of domain knowledge. Following the ontology development framework proposed in (Holsapple at al., 2002), we aimed to support domain experts to reach consensus through iterative evaluations and improvements of an initial ontology.

Before starting designing the initial ontology, we did an extensive literature review and discussed with domain experts about the scope of the top-level ontology. Domain experts were carefully selected in order to complement each other and represent diverse viewpoints resulting to a group of five academics and five practitioners with extensive experience in VOs.

To design the initial ontology we used the ECOLEAD top-level ontology as a starting point for our work. We then utilized ontology learning tools to analyze a corpus of 79 papers from the related literature with the aim to identify important terms and relationships between terms. This process has been leveraged with Text2Onto (Cimiano et al., 2005), an ontology learning tool. We then identified the terms of the ontology and derived class definitions and class hierarchy. We followed the top-down approach and took into account suggestions for class hierarchies provided by Text2Onto. Next, we determined the properties of classes; suggestions for object properties from Text2Onto were again taken into account. Finally, we determined the restrictions of the data type and the object properties. Having the initial ontology at hand, we worked with experts to evolve the initial version by asking them to evaluate it and finally reach consensus and agree upon the final version. To reach consensus between experts that were not co-located and did not collaborate synchronously, we followed an adaptation of the Delphi method (Fitch et al., 2001), a technique which involves multiple iterative rounds of anonymous responses to a questionnaire until either the opinions converge or until no further substantial change in the opinions can be elicited. In each round, participants were asked to rank using a 5-point Likert scale each concept, and each taxonomic and non-taxonomic relation of concepts for relevancy to the project and for ambiguity. Moreover, for each concept synonyms were collected in order to broaden the vocabulary of the domain. Finally, participants could enter new concepts and relations in each round which were then fed again into the evaluation process. The facilitator provided details about particular items for which no consensus was reached and participants rated them again. The iterative process continued until all participants agreed on all items.

4   THE OCEAN ONTOLOGY

The Ocean ontology aims to represent a conceptual schema of the domain of VOs typically referred to as terminology box or TBox. The domain of VOs includes concepts such as collaborative network organization, virtual breeding environment and business opportunity that model the external environment in which VOs are being bred; such concepts are modeled in the ECOLEAD ontology. OCEAN mainly focuses on knowledge-oriented collaborations apposite for VOs. Nevertheless, to fully cover the domain of VOs, we have used the part of the ECOLEAD ontology which covers extensively the VO breeding environment and built upon it towards a unified model that captures the general aspects of collaborative network organizations and at the same time present details about knowledge-oriented collaborations that are important during the creation, operation and termination phases of VOs .

For developing the OCEAN ontology we have used Protégé (Protégé) and for validating it we have used the OWL-DL reasoner Pellet (Pellet). Pellet provides reasoning services and performs consistency checking and computation of inferred hierarchies, equivalent classes and inferred individual types (Sirin et al., 2007). Due to spatial restrictions we can not depict the whole (53 terms and 77 relationships were identified and modeled) of OCEAN; instead we depict the critical concepts, only. The OWL-DL representation of the complete OCEAN top level ontology is available online at: http://www.imu.iccs.gr/ontologies/ocean/. We have categorized the critical OCEAN concepts into: Breeding Environment related OCEAN concepts and
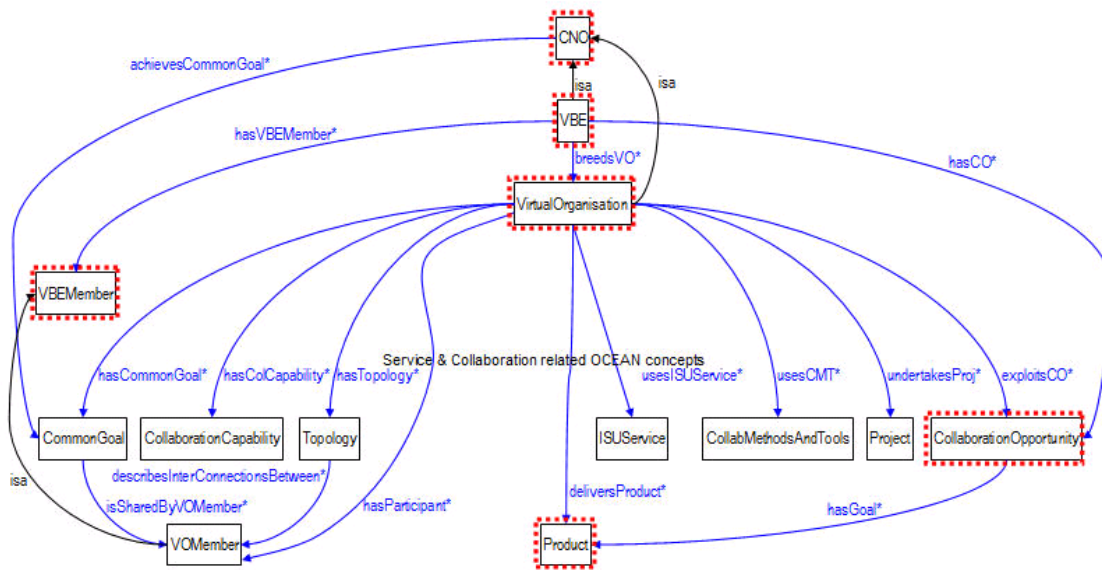
Figure 1: Breeding Environment related OCEAN concepts.

Service & Collaboration related OCEAN concepts (refer to knowledge-enabled collaboration services).

4.1 Breeding Environment Related OCEAN Concepts

We have organised the presentation of OCEAN by putting first concepts and relationships that describe the VO's breeding environment, as a necessary artifact to describe the full picture of the domain (figure 1). The highlighted concepts were taken from the ECOLEAD ontology, while the remaining concepts and relationships appear as extensions. Some of the breeding environment related OCEAN concepts and relationships are:

A virtual organization is a short-term association (of organizations) with a specific goal of acquiring and fulfilling a collaboration opportunity. A VOmember represents an entity collaborating with other entities in the VO (Plisson et al., 2007). In simpler words VOmembers are the organizations which participate in a VO. A virtual organization is bread in a VBE, an association of organizations and their related supporting institutions, which have both the potential and the will to cooperate with each other through the establishment of a base long-term cooperation agreement and interoperable infrastructure (Camarinha-Matos et al., 2005). VO's aim is to deliver Products (anything an organization may produce: goods or services), has a CommonGoal, undertakes a Project, uses CollaborativeMethodsAndTools and exploits a

CollaborationOpportunity. With the term CollaborativeMethodsAndTools we define all the synchronous or asynchronous tools and methods that are going to be developed in terms of a system to support and enhance collaboration within a VO.

Every VOmember has (or should have) CollaborationCapability which declares the capability that is relevant to the participation of an enterprise in collaboration with partner enterprises. It includes both HR capabilities of personnel involved in management and operation of collaborative activities, and interoperability of software systems. The concept of CollaborationCapability concerns mainly the pre-creation phase of a VO (i.e. identification phase for (Plisson et al., 2007)) as it focuses on the knowledge about the capability of future VO partners to collaborate. A critical factor, that is often disregarded in efforts that describe and support VOs, is the fact that two potential partners may be unable to collaborate, although they appear to have all the necessary assets for participating in a specific VO (e.g. two partners that had unsuccessful collaborations in previous VOs, partners that have been engaged in lawsuits against each other etc.). Within the system that will use the OCEAN top level ontology, a VO may use an ISUService (described in the next section).

The structure of a VO is described with the term topology which stands for the arrangement of the participants inside the VO (e.g. Star Alliance: A grouping of independent organizations, with a core organization taking the lead). By declaring that a VO

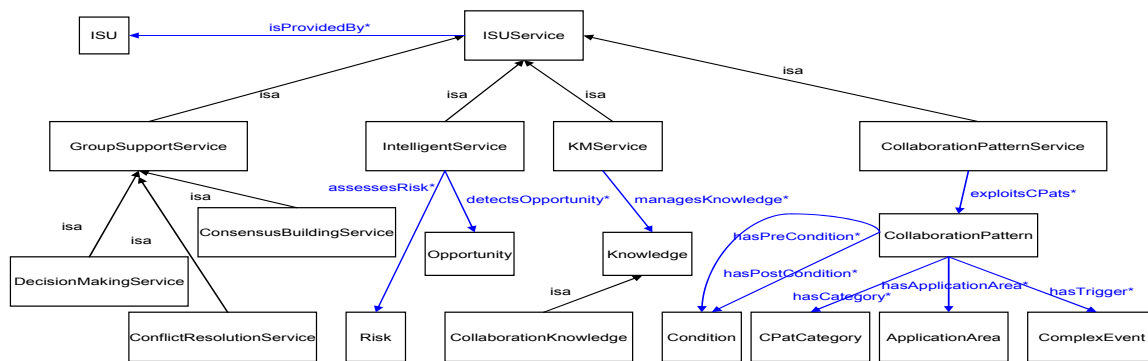Figure 2: Service & Collaboration related OCEAN concepts.

is a kind of CNO we express that a VO is a collaborative network of organizations.

4.2 Service & Collaboration Related OCEAN Concepts

In this section we present the top-level ontology concepts that refer to collaborations and services (figure 2) that are to be provided by the ISU. The Interoperability Service Utility (ISU) is the enabling system of services for delivering basic interoperability to enterprises, independent of particular IT deployment. It may also denote and an enterprise providing such services. A service is a provider-client interaction that creates and captures value (IBM). An ISU service is technical, commoditized functionality, delivered as services provided by an ISU to support the collaboration between enterprises. A non-exhaustive list of ISU services is presented below. Lower-level domain ontologies further specify each one of the ISU services.

DecisionMaking, ConsensusBuilding, ConflictResolution services and other Group Support Services. For example, reach decision on production plans, budget expenditure, etc.

KnowledgeManagementServices helping a company that wants to enter the VO, to efficiently build up and manage a knowledge base of collaboration-oriented internal knowledge, together with knowledge sharing and exchange services which guarantee adequate treatment of confidentiality concerns.

Specific IntelligentServices such as OpportunityDetection (e.g., detection of opportunity to develop a new product) and RiskAssessment (e.g., risk of failure of the new product, risk of conflict between partners).

CollaborationPatternServices as a means to use and reuse proven, useful, experience-based ways of

doing and organizing communication and collaboration activities in specific knowledge-oriented collaborative tasks. A Collaboration Pattern has Pre-Conditions, Post-Conditions, category (CPatCategory), Application Area, and Triggers that are comprised of Complex Events.

5 APPLICATION OF THE OCEAN ONTOLOGY

In this section, we present the application of the OCEAN ontology and architecture for network enterprise collaboration in the pharmaceutical industry. The pharmaceutical industry is considered a typical example of knowledge-intensive sector where the problem of dealing with heterogeneous and vast number of information appears to be insurmountable.

According to Investigational New Drug Application Process (IND), the process of developing a new dermatological drug involves several different stages starting from pre-clinic studies (testing the drug in the lab, use it on guinea pigs etc.) and continuing with the four phases imposed by Foods and Drug Administration (FDA) and the European Medicines Agency (EMEA). During these phases a formal proposal is introduced to the FDA or EMEA with all the details of the new drug. Upon approval, phase one starts with the testing on a group of healthy people in order to decide on the drug toxicity, liver and spleen reaction, the best dose amount, the best way to administer the new drug (oral, patch, intravenous, intradermal). The next two phases involve the testing on a group of sick people in order to decide on the new dermatological drug effectiveness. Phase two involves 100-300 sick people while phase three involves the testing on an extensive group with ethnographic differences that takes place in different
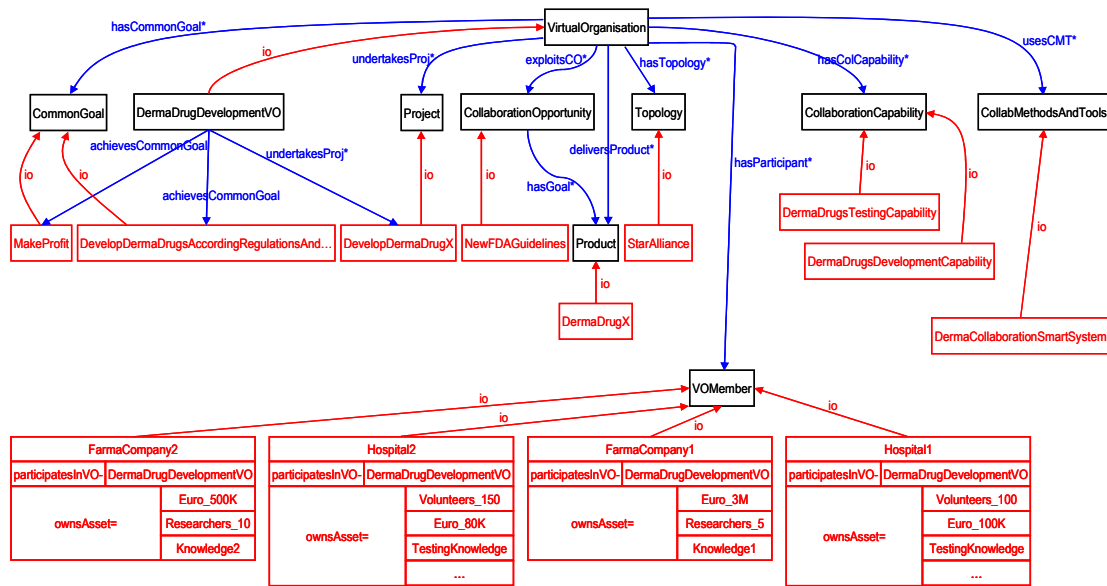
Figure 3: Breeding Environment related OCEAN concepts – Instantiated.

hospitals. Since only a 5% of new drugs are approved to be circulated in the public, not many efforts continue with Phase four where the approved drugs continue to be tested for side effects for many years after their first circulation.

In our case, we consider that the new dermatological drug has reached the critical phase three where the testing must proceed in different hospitals. According to the ICH (International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use) (ICH) that was held in Helsinki four decades ago, there was an agreement upon a set of good clinical practices. Of course these best practices may be altered by the ethics committees of each country involved that may decide on the details of the drug testing (e.g. people with age less than fourteen should not be tested) or by the release of a new regulation from the FDA or EMEA. Such a change on the clinical practices can be considered as a new opportunity in terms of a VO.

As shown in figure 3, the OCEAN ontology has been instantiated in order to describe our domain. The VO follows a certain topology: Star Alliance. This specific topology for structuring a VO involves the grouping of independent organizations, with a core organization taking the lead management role.

The VO comprises two pharmaceutical companies with expertise in dermatological drug development and two hospitals with their own assets (testing knowledge, doctors supervising and volunteers). The common goal for this VO has been agreed to be the development of dermatological

drugs according to the regulations and ethics taking into account the profit maximization. The VO has been bred by a drug development virtual breeding environment (VBE) that combines pharmaceutical companies that are capable of developing any new drug and hospitals for the testing processes.

5.1 Enabling shared Understanding

The ability of OCEAN to provide a common terminological reference and a shared understanding for human participating in VOs, is demonstrated by the following set of questions for which we were able to get answers from our instantiated ontology. We have used the SPARQL language for assessing the expressiveness capability of OCEAN. SPARQL is a query language for the Semantic Web that can be used to query an RDF Schema or OWL model in order to filter out individuals with specific characteristics (SPARQL).
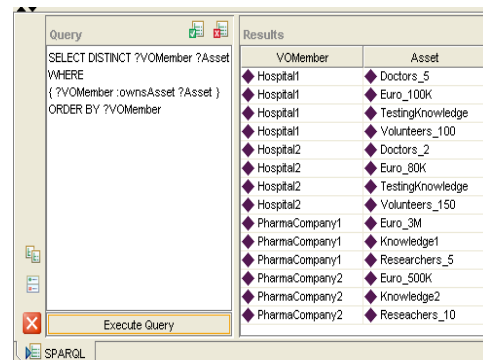
Figure 4: Retrieval of VO members' assets.

One such question could be: Which are the assets of each VO member? In figure 4, it is shown how we can make such a question using SPARQL. Regarding our application, we get as an answer the group of assets per VO members (Hospital 1, Hospital 2, Pharma Company 1 & 2).

In table 3 the reader can find more questions that can be answered using SPARQL queries through the instantiated OCEAN top level ontology.

Table 3: SPARQL Queries.

| QUERY	SPARQL QUERY
In which VOs have a specific pharmaceutical company participated in the past?	**SELECT ?VO** **WHERE {** **:PharmaCompany1 :particapatesInVO ?VO .** **}**
Which are the projects that the DermaDrugDevelpmentVO has undertaken so far?	**SELECT ?Proj** **WHERE {** **:DermaDrugDevelopmentVO :undertakesProj ?Proj .** **}**
What are the possible moderator services depending on common goals for the DermaDrugDevelpmentVO?	**SELECT DISTINCT ?ModSrv ?CGoal** **WHERE {** **:DermaDrugDevelopmentVO :hasCommonGoal ?CGoal .** **:DermaDrugDevelopmentVO :usesModeratorService ?ModSrv** **}**

Unlike databases, ontologies built in OWL such as OCEAN has a so-called open-world semantics in which missing information is treated as unknown rather than as false and OWL axioms behave like inference rules rather than as database constraints. For example, if we have asserted that BiotechOne is a VO Member and that it Participates In (which is the inverse property of hasParticipant) BioAlliance, then, because only Virtual Organizations have VO Members as participants, this leads to the implication that BioAlliance is a Virtual Organization. If we were to query the ontology for instances of Virtual Organization, then BioAlliance would be part of the answer. We can also ask if any Collaborative Network Organization that has VO Members as Participants is necessarily of Virtual Organization. Query answering in OWL is analogous to theorem proving; therefore the OCEAN top level ontology plays itself an important role and is actively considered at query time. Considering both the schema and the data represented in OCEAN can be very powerful, making it possible to answer conceptual and extensional, queries as well as to deal with incomplete information.

6 CONCLUSIONS

In this paper we presented OCEAN, a top-level ontology for collaborative networked organizations. The OCEAN ontology covers the creation, operation and termination phases of VOs and is designed as a lightweight top-level ontology that provides a common terminological reference for VO concepts and relations. We validated the OCEAN ontology as an expressive tool for describing such VOs using SPARQL queries.

We believe that the OCEAN ontology formalizes and enables network enterprise collaboration as it models formally the main factors that affect/enable the network enterprise collaboration orchestrated by an entire system. It targets specifically the relationships between "high level pieces" of domain knowledge, explaining how they contribute altogether to the network enterprise collaboration. This top level ontology also enables better communication by defining a common-agreed vocabulary that: ensures shared meaning and understanding regarding project goals; facilitates knowledge acquisition in situations where teams have to work together because the ontology becomes a common, agreed-upon understanding of the terms, which can be understood by team members with different background knowledge (Valente et al., 1996). Ultimately, the OCEAN ontology supports semantic interoperability between software components by formalizing the used vocabulary explicitly in a machine-readable form. This is possible due to the openness of the OCEAN top level ontology which will act as "glue" between other domain ontologies that describe specifics of any VO, VO member, knowledge related functionalities and assets. Although, we briefly described here the application of the OCEAN ontology in the pharmaceutical sector, we intend to also use it in the manufacturing industry, in terms of the SYNERGY ICT project for considering its applicability and address possible limitations with appropriate extensions of the top-level ontology.

ACKNOWLEDGEMENTS

REFERENCES

B. Andersson, M. Bergholtz, A. Edirisuriya, T. Ilayperuma, P. Johannesson, B. Grégoire, M. Schmitt, E. Dubois, S. Abels, A. Hahn, J. Gordijn, H. Weigand, B. Wangler, Towards a Reference Ontology for Business Models, in: *Proceedings of the 25th International Conference on Conceptual Modeling (ER2006)*, Tucson, Arizona, USA, 2006.

T. Berneers-Lee, Welcome to the Semantic Web, The Economist, 21 (2007) 146. *Business Management Ontology*, available online at: http://www.bpiresearch.com/Resources/RE_OSSOnt/re_ossont.htm

L. M. Camarinha-Matos, H. Afsarmanesh, Collaborative networks: A new scientific discipline, *Journal of Intelligent Manufacturing* 16(2005) 439–452.

S. Castano, A. Ferrara, S. Montanelli. Ontology-based Interoperability Services for Semantic Collaboration in Open Networked Systems. *Springer, Interoperability of Enterprise Software and Applications*, ISBN: 978-1-84628-151-8, 2006

C. M. Chituc, C. Toscano, A. Azevedo, Interoperability in Collaborative Networks: Independent and industry-specific initiatives – The case of the footwear industry, *Computers in Industry* 59 (2008) 741–757.

P. Cimiano, J. VÄolker, Text2Onto A Framework for Ontology Learning and Data-driven Change Discovery in: *Proceedings of the 10th International Conference on Applications of Natural Language to Information Systems*, 2005.

M. Cristani, R. Cuel, A survey on ontology creation methodologies, *International Journal on Semantic Web & Information Systems 1* (2) (2005) 49–69.

W. H. Davidow, M. S. Malone, The Virtual Corporation, Harper-Collins (Eds.), New York, 1992.

DIP Business Data Ontology, available online at: http://dip.semanticweb.org

ebXML, available online at: http://www.ebxml.org

EDI, UN/EDIFACT. available online at: http://www.unece.org/trade/untdid

European Medicines Agency (EMEA), available online at: http://www.emea.europa.eu/

C. Fellbaum, WordNet, An Electronic Lexical Database, *MIT Press*, Cambridge, MA, 1998.

K. Fitch, S. J. Bernstein, M. S. Aguilar, B. Burnand, J. R. LaCalle, P. Lazaro, M. van het Loo, J. McDonnell, J. Vader, J. P. Kahan, The RAND/UCLA Appropriateness Method User's Manual, 2001. Foods and Drug Administration (FDA), available online at: http://www.fda.gov/

A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, L. Schneider, Sweetening Ontologies with DOLCE, in: A. Gomez-Perez and V. R. Benjamins (Eds.) *Knowledge Engineering and Knowledge*

C. W. Holsapple, K. D. Joshi, Collaborative Approach in Ontology Design, in: *Proceedings of the Communications of the ACM* 45 (2002).

N. Huang, S. Diao, Ontology based enterprise knowledge integration, Robotics and Computer Integrated Manufacturing, 2007.

IBM Services Sciences, Management and Engineering, available online at: http://www.research.ibm.com/ssme/

ICH, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, available online at: http://www.ich.org

Investigational New Drug Application Process (IND), available online at: http://www.fda.gov/cder/Regulatory/applications/ind_page_1.htm

A. Kiryakov, Ontologies for Knowledge Management, in: Davies, J., Studer, R., Warren, P. (Eds) *Semantic Web Technologies: Trends and Research in Ontology-based Systems*, Wiley, 2006.

H. K. Lin, J. A. Harding, A manufacturing system engineering ontology model on the Semantic web for inter-enterprise collaboration, Computers in Industry 58 (2007) 428–437.

A. Mowshowitz, Virtual Organization, in: *Proceedings of Communications of the ACM* 40 (1997) 30–37.

I. Niles, A. Pease, Towards a standard upper ontology, in: Proceedings of the International *Conference on Formal ontology in information Systems* (FOIS '01, ACM), 2-9, USA, 2001.

PELLET, OWL-DL ontology reasoner, available online at: http://clarkparsia.com/pellet/

J. Plisson, P. Ljubic, I. Mozetic, N. Lavrac, An ontology for virtual organisation breeding environments, *IEEE Transactions on Systems, Man, and Cybernetics* – Part C: Applications and Reviews, 37(6): 1327-1341, 2007.

Protégé, Ontology Development Platform, available online at: http://protege.stanford.edu/

D. Rajpathak, M. Zdenek, R. Roy, A Generic Library of Problem Solving Methods for Scheduling Applications, *IEEE Transactions on Knowledge Data Engineering* 18(6): 815-828, 2006.

RosetaNet, available online at: http://www.rosettanet.org

T. Ruokolainen, Y. Naudet, T. Latour. An Ontology of Interoperability in Inter-Enterprise Communities. *Springer, Enterprise Interoperability II*, ISBN: 978-1-84628-857-9, 2007

E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur, Y. Katz. Pellet: A practical OWL-DL reasoner in: *Proceedings of the Web Semantics: Science, Services and Agents on the World Wide Web* 5(2): 51-53, 2007.

SPARQL, ontology query language, available online at: http://www.w3.org/TR/rdf-sparql-query/

TOVE ontologies, available online at: http://www.eil.utoronto.ca/tove/ontoTOC.html

M. Uschold, M. King, S. Moralee, Y. Zorgios, Y. The enterprise ontology, Knowledge Engineering Review 13, 1998.

Y. Verginadis, D. Apostolou, N. Papageorgiou, G. Mentzas. An Architecture for Collaboration Patterns in Agile Event-Driven Environments. *Fourth IEEE Workshop on Agile Cooperative Process-Aware Information Systems* (ProGility 2009), June 29 - July 1, 2009, Groningen

A. Valente, J.A.P.J Breuker. Towards principled core ontologies. SRDG Publications, University of Calgary, 1996, available online at: http://dare.uva.nl/record/15931

POSTER

USERS' PERSONALITY TRAITS IN THE CONTEXT OF VIRTUAL REALITY

Ioannis G. Krasonikolakis, Adam P. Vrechopoulos and Athanasia Pouloudi

Department of Management Science and Technology, Athens University of Economics and Business
76 Patission Str., 10434 Athens, Greece
{krasos, avrehop, pouloudi}@aueb.gr

Abstract: The purpose of using the internet has been thoroughly studied in the context of "traditional" web. This exploratory study aims to illustrate the profile of users visiting Virtual Worlds (VWs) through the Web. The findings confirm VWs' social origin and highlight entertainment as one of the most attractive purposes of entering to this virtual world. The study also contributes to our understanding of virtual reality retailing dynamics and sets the further research agenda.

1 INTRODUCTION

In Virtual Worlds (VWs), users from all over the world can interact within a rich multimedia three-dimensional environment. Participants in these worlds can engage in socialization, information search, entertainment, education, and e-commerce activities. In order to better understand user behaviour in the context of VWs, the objective of the present study is to profile user behaviour in this emerging electronic environment through an empirical invesigation of user preferences.

2 LITERATURE REVIEW

2.1 Virtual Reality Dynamics

Social Networking has been the precursor of VWs (Messinger, Stroulia, Lyons, Bone, Niu, Smirnov, and Perelgut, 2009). The first social networking application has been SixDegrees.com that was launched in 1997 (Boyd and Ellison, 2007). Thereinafter, numerous other social web sites and applications followed, such as Twitter, MSN, YouTube and Facebook, adopting several Web 2.0 or 2.5 applications. No matter what their orientation is, the social aspect has been the common denominator.

These platforms are technology enabled to facilitate various business activities and opportunities. Virtual teams can work together in the development of new products, attend virtual meetings from their desk, test new products in a manipulated environment under low risk, or be trained to a new position. The flexibility of collaboration coupled with the highly vivid interfaces help them mimic real world activities.

2.2 User/Consumer Behaviour

The common interests of the members of a virtual community form and shape that community (Porter, 2004). Similarly, in the virtual reality context, there are VWs such as Second Life that induce users to be part of activities and behaviour that are unusual or do not exist in real life, and other that support members to retain and increase real life behaviour and social activities (Messinger et al. 2009).

According to Bellman, Lohse and Johnson (1999, p.37), "the most important information for predicting online shopping habits are measures of past behaviour". Along these lines, O'Keefe, Cole, Chau, Massey, Montoya-Weiss and Perry (2000) demonstrated that there are considerable differences in the purpose of using the Internet between different groups of subjects. In order to measure the "Purpose of Internet Use", they used four constructs; social communication, E-Commerce, information search and hobby. These constructs form the basis for our empirical investigation in the present paper.

3 RESEARCH OBJECTIVES AND METHODOLOGY

An exploratory quantitative research was considered as the most appropriate approach to address the aforementioned goals. To that end, an electronic questionnaire was developed and served as the data collection instrument of the study. Data were collected through groups on Facebook that are funs of VWs and from questionnaires distributed in Second Life. 104 valid responses were collected. (See Krasonikolakis, Vrechopoulos and Pouloudi (2010) for further details on the research methodology of the present research).

Employing the "Purpose of Internet Use" construct discussed in section 2 (O' Keefe et al. 2000), it was first attempted to segment the sample according to the construct that each subject classified to. The sample was segmented into three groups. Specifically, the first group, labelled "Social Communication", involves users that visit VWs to satisfy their social needs only. The second group labelled "E-Commerce" involves users that visit VWs for e-commerce purposes regardless of any other activities they are engaged in. Finally, the third group involves the remaining users that visit VWs for all purposes except e-commerce. The relevant answers and corresponding segmentation of the sample are summarized in Table 1.

Table 1: Segmentation of groups according to subjects' answers.

Question: What do you usually do in a virtual world?	
Group	Answer
Group #1: "Social Communication" (only)	Meet friends
	Meet new people
Group 2: "E-Commerce" (and other activities)	Shopping
	Sell goods
Group #3: "Hobbies" and/or "Information Search" and/or "Social Communication"	Dancing (Hobby)
	Play games (Hobby)
	Build things (i.e. home furniture, car etc) (Hobby)
	Information search
	Education (Information Search)
	Other

4 ANALYSIS OF RESULTS

4.1 Demographics

The gender dimension of the participants was split roughly evenly (53,8% being male), while most of the respondents were under 35 years old (81,8%).

Table 2: Demographic characteristics of the sample.

Demographics		N=104 (%)
Age	<18	2,9
	18-25	40,4
	26-35	38,5
	36-50	16,3
	>50	1,9
Gender	Male	53,8
	Female	46,2
Education	High School	19,2
	Undergraduate	26,9
	Graduate	31,8
	Postgraduate	13,5
	PhD	7,7
Nationality	Greek	87,5
	European	10,6
	Australian	1
	American	1
Average income in Euros	0-700	17,3
	701-1100	21,2
	1101-1500	27
	1501-3000	16,3
	>3000	1
	N/A	17,3

At least 65,5% of the population have an average income up to 1500 Euros whereas the majority of the respondents were Greek (87,4%).

4.2 Internet Behaviour

The first step for mapping user behaviour has been to record the frequency that each group visited the internet (Table 3). About 46,7% of the respondents that belong in the "Social Communication" group use the internet many times or at least once per day. The corresponding percentage is much greater (82,5%) for "E-Commerce" users (i.e. group #2) and about the same (41,2%) for the third group.

Table 3: Frequency of using the internet.

Frequency(%)	Group #1	Group #2	Group #3
Many times per day	20	30	26,5
At least every day	26,7	52,5	14,7
Sometimes per week	20	12,5	23,5
At least once a week	23,3	5	23,5
Sometimes per month	10	0	11,8

It is probable that someone buys or sells products over the internet but not in a VW and vice versa. The results (Table 4) indicate, however, that a predictable 97,5% of "E-Commerce" users buy

products over the internet. A significant amount (53,3%) of group #1 users buy products over the internet while they are not engaged in shopping activities in VWs. Finally, the same applies to the subjects of the third group (58,9% of them buy products through the Internet).

Table 4: Consuming aspect (buying products through the internet).

Consuming aspect (%)	Group #1	Group #2	Group #3
Yes	53,3	97,5	58,9
No	46,7	2,5	41,1

The frequency of buying products over the internet is summarized in Table 5. Approximately half of the users (50,1%) that visit VWs only for communication purposes, buy products over the internet at least once or twice a month.

Table 5: Frequency of buying products over the internet (responses as a percentage of consumers of the respective groups).

Frequency (%)	Group #1	Group #2	Group #3
Almost every week	18,8	20,5	15
Once or twice a month (approximately)	31,3	41	25
Once or twice in six months (approximately)	37,5	25,6	30
Once or twice a year (approximately)	12,5	7,7	20
Once or twice in the past	0	5,1	10

The corresponding proportion for "E-Commerce" users is greater (61,5%) and for the third group is 40%.

Table 6: Kind of products bought over the internet (responses as a percentage of consumers of the respective groups).

Kind of products(%)	Group #1	Group #2	Group #3
Only tangible	25	38,5	40
Only intangible	6,3	10,3	15
Both tangible and intangible	68,8	51,3	45

Finally (Table 6), regarding the nature of products that users buy, the proportions are lower in intangible products in all groups. Indicatively, only 6,3% of "Social Communication" users buy only

intangible products and the percentages for the other groups are 10,3% and 15%, respectively.

4.3 User Behaviour in Virtual Worlds

According to the results presented in Table 7, 83,3% of the "Social Communication" users (group 1) visit VWs at least once a week. The corresponding percentage is greater (92,5%) for "E-Commerce" users and for the third group (76,5%).

Table 7: Frequency of visiting virtual worlds.

Frequency(%)	Group #1	Group #2	Group #3
Never	0	0	5,9
Every day	30	47,5	26,5
Every two days	13,3	37,5	23,5
Once a week	40	7,5	26,5
Once a month	10	7,5	14,7
Once a year	6,7	0	2,9

The following Table (Table 8) highlights the social aspect of VWs. It is noteworthy that the percentage of users of group #1 (13,3%) and #3 (5,9%) that do not visit other social web sites is greater than that of "E-Commerce" users (2,5%).

Table 8: Visit of other social web sites and applications such as Facebook, MSN, MySpace, etc.

Visit of other social Web sites(%)	Group #1	Group #2	Group #3
Yes	86,7	97,5	94,1
No	13,3	2,5	5,9

As part of our study of VW user profiles, we also investigated how the users first learned about the existence of VWs (Table 9). For the first group, most of the users (86,8%) learned about VWs from friends (offline and online) and through e-mails. The same applies to 75% of the respondents of the second group and 70,6% of the third group. It is notable that only 5,9% of the respondents of the third group were informed through scientific articles and journals, while 20% of "E-Commerce" users, randomly.

Table 9: Learning about VWs.

Frequency(%)	Group #1	Group #2	Group #3
friends offline	33,4	20	23,5
friends online	26,7	42,5	32,4
advertisements	3,3	2,5	5,9
e-mail	26,7	12,5	14,7
scientific articles or journals	0	2,5	5,9
randomly	10	20	17,7

The majority of users, especially those of the first two groups seem to embrace the idea that VWs are becoming an emerging alternative retail channel. Nevertheless, approximately one in four (26,5%) of the users of the third group do not (Table 10).

Table 10: Consideration of VWs as an emerging alternative retail channel.

Emerging alternative retail channel(%)	Group #1	Group #2	Group #3
Yes	86,7	92,5	73,5
No	13,3	7,5	26,5

Looking further into the perception of VWs as an e-business outlet, we investigated what types of stores or business users visit in VWs. As users had the ability of choosing more than one option, Table 11 depicts the percentages of users that chose only one option and the percentage of users that chose more than one option (combination). The findings show that 33,3%, 15% and 20,1% of the users within each group respectively (i.e. for groups 1,2 and 3), visit apparel stores only. However, the frequency that the second "E-Commerce" users group visit a combination of the stores, is greater (77,5%) to that of the first "Social Communication" users group.

Table 11: Types of stores/business visited in a VW.

Types of stores/ businesses (%)	Group #1	Group #2	Group #3
Apparel	33,3	15	20,1
Hotels	10	0	11,8
Furniture (Home equipment)	6,7	0	2,9
Consulting services for consumers	6,7	5	5,9
Grocery	0	0	0
Non-profit organisations	3,4	2,5	8,8
Combination of the above	39,9	77,5	50,5

5 CONCLUSIONS

An important finding of the present study is the great amount of users that conduct e-commerce transactions in the "traditional" Web but do not buy products over the internet in the context of VWs. While this merits further exploration, we posit that it can probably be explained either because users treat VWs as an entertaining or gaming oriented environment and not as a retailing channel, or because they are considered light users of VWs and are reluctant to commit to transactions in an environment that is deemed unstable.

The origins of VWs in social computing (cf. Messinger et al. 2009; Chittaro and Ranon, 2002) has also been confirmed in the present study: a great percentage (28.8%) of the sample visit VWs only to meet friends or meet new people (social aspect). This was also confirmed by the finding that a great amount of users of all groups (i.e. 86,7%, 97,5%, and 94.1%, respectively) also visit other social networks (i.e. YouTube, Facebook, MSN etc). Therefore, it is important to take this consideration into account when investigating user/consumer behaviour in V-Commerce.

In conclusion, the present study contributes to our understanding of the purpose of VWs use by empirically examining the user behavioural and demographic patterns in the virtual reality context. It demonstrates that both the social aspect and past experience play a significant role in users' "virtual" decisions and behavioural habits.

REFERENCES

Bellman, S., Lohse, G.L. and Johnson, E.J. (1999). Predictors of Online Buying Behaviour. *Communications of the ACM,* 12, 32-37.

Boyd, Ellison, N. (2007). Social network sites: definition, history, and scholarship. *J. of Computer-Mediated Communication,* 13 (1), Article 11.

Chittaro, L., Ranon, R. (2002). New directions for the design of virtual reality interfaces to E-Commerce sites. *5th Int. Conf. on Advanced Visual Interfaces.* New York: ACM Press.

Krasonikolakis, I., Vrechopoulos, A., Pouloudi, A., (2010). Users' habits in virtual worlds, *ELTRUN HERMES Newsletter,* Issue No. 61 (April – May 2010).

Messinger, P., Stroulia, E., Lyons, K., Bone, M., Niu, R., Smirnov, K., Perelgut, S. (2009). Virtual Worlds – past, present, and future: New directions in social computing, *Decision Support Systems,* 47, 204–228.

O'Keefe, R., Cole, M., Chau, P.Y.K., Massey, A., Montoya-Weiss, M. and Perry, M. (2000). From the User Interface to the Consumer Interface: Results from a Global Experiment. *International J. of Human Computer Studies,* 53(4), 611-628.

Porter, C., (2004). A typology of virtual communities: a multi-disciplinary foundation for future research, *J. of Computer-Mediated Communication,* 10 (1) Article 3.

TECHNOLOGY PLATFORMS

SHORT PAPERS

MIDDLEWARE INTEGRATION PLATFORMS: A NEW CHALLENGE TO BUSINESS MODELS OF ICT COMPANIES
Unleashing the Business Potential of Horizontalization

Markus Eurich

Technology and Innovation Management (TIM), Swiss Federal Institute of Technology Zürich (ETH)
Scheuchzerstrasse 7, 8092 Zürich, Switzerland
SAP (Switzerland) Inc. Research, Kreuzplatz 20, 8008 Zürich, Switzerland
meurich@ethz.ch

Roman Boutellier

TIM, Department for Management, Technology, and Economics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland
roman.boutellier@sl.ethz.ch

Abstract: Information and Communication Technology (ICT) has the potential to transform business on different levels. Changes in the ICT market, the technology and its dynamics, and of customers' demands and needs will challenge the current business models of ICT companies. If middleware integration platforms integrate presently isolated networks, organizations may have to re-define their role and their business scope; especially if the assumptions hold true that middleware platforms foster disintermediation and enable new players to enter the market. In this paper, the business potential of reusing sensing, actuation and processing services for different applications - termed horizontalization - is described. It provides insights to the design and architecture of middleware integration platforms designed to enable horizontal services and gives examples of e-business opportunities that can be realized on top of middleware integration platforms.

1 INTRODUCTION

In the early stages of the Internet it was primarily considered as an additional information channel. It took some time before companies realized the potential of the Internet to increase efficiency: the Internet enables transformation of business processes and business networks, and eventually even business scope redefinitions (Venkatraman, 1994). The evolution of the Internet has still not come to an end and latest advances in Information and Communication Technology (ICT) extend the Internet's business potential. One of the next evolutionary steps is assumed to be the bridging of the physical world with the digital one by the means of recent innovations in the fields of (Wireless) Sensor and Actuator Networks (WSAN), Radio Frequency Identification (RFID), Machine to Machine (M2M) data communications, positioning systems, and Network-Embedded Devices (NED) (Presser et al., 2008). Vertically integrated and

closed solutions on the basis of WSANs, RFID and M2M are already deployed for various purposes. Middleware integration platforms now aim at integrating information from different sources, which means that they approach to combine vertical solutions that were originally designed for a single purpose. In comparison to so-called vertical solutions (built for a single purpose), we term the concept of reusing services in a uniform manner - such as sensing, actuation and processing services - *horizontalization*. In this way, the speed of business transformation could increase, as the reach of the Internet is extended to the real world and its potential can be realized on a global inter-organizational level. Information is not just an auxiliary factor that supports the production and shipment of physical products; it is even a product of its own in electronic platforms that enable the processing of information, communications and transactions. The evolution of the Internet may enable new e-business opportunities and new business models, which can generate new value. E-

239

business opportunities can be of different nature, e.g. Business-to-Business (B2B), Business-to-Consumer (B2C), administration to business, or intra-organizational (e.g., Amit and Zott, 2001; Timmers, 1998); this paper focuses on B2B and B2C.

The purpose of this paper is to foster the understanding of the business potential of horizontalization. Therefore, future directions of the Internet and a prime example of a middleware integration platform that aims at realizing the concept of horizontalization are described in section 3. Concrete e-business opportunities on top of such a middleware integration platform are depicted in section 4. Section 5 is dedicated to implications of the concept of horizontalization and the deployment of middleware integration platforms. These implications suggest a challenge to the business models of ICT companies and may help managers and strategist to anticipate a redesign of business processes and networks and eventually even the overall business scope.

2 METHODOLOGY

The research questions were: (1) how should a middleware integration platform be designed and how does it work? (2) Which e-business opportunities may emerge on top of a middleware integration platform? (3) What implications on business models arise from the deployment of middleware integration platforms and especially the concept of horizontalization?

As we found little research has been done that combines the aspects, in which we want to investigate, and a comprehensive understanding of those aspects seems missing in the current literature, a qualitative research method was applied (on basis of Bryman and Bell, 2007). The research approach was divided into two parts: the design of middleware integration platforms and the assessment of e-business opportunities that may emerge on top of those platforms were investigated within the framework of the research project SENSEI ("Integrating the physical with the digital world of the network of the future", SENSEI, 2010a). We conducted 25 interviews to assess design principles and business opportunities of middleware integration platforms. The interviewed organizations were ICT businesses, to which the SENSEI approach may become relevant; this particularly includes: infrastructure providers, service developers, and service providers. We identified key personnel in charge of future strategy, experts in Future Internet and WSAN research as the valuable contact persons

and potential interviewees. The expert interviews have been conducted via phone or face-to-face and followed an interview guideline. The other part of our research approach went beyond the SENSEI project: a literature review was performed, which was followed by another round of in-depth expert interviews. Eight semi-structured face-to-face and phone interviews were conducted with managers and strategists to foster the understanding of how middleware integration platforms and especially the concept of horizontalization would challenge their business models.

3 MIDDLEWARE PLATFORMS

This section describes related work and the results of a literature review about future directions of e-business and it presents the SENSEI framework as a prime example of middleware integration platforms. We validated SENSEI's design by 25 in-depth expert interviews.

As the Internet evolves e-business services can become more and more sophisticated. One of the most important drivers of the Internet's evolution is its extension of reach to the real world: innovations that aim at connecting the physical with the digital world in a network of objects are commonly termed the "Internet of Things" (IoT; cf. e.g. Floerkemeier et al., 2008). ReadWriteWeb (2009) provides a good overview of recent IoT products. Recent ideas picture the IoT as an "encompassing vision to integrate the real world into the Internet" (Presser et al., 2008): also referred to as the Real World Internet (RWI). The RWI dimension and other concepts such as the Internet-of-Services (Buxmann et al., 2009) are parts of the so-called "Future Internet" (European Future Internet Portal, 2010; Presser et al., 2008). The evolution of the Internet and especially the RWI dimension of the Future Internet are only possible due to advances in ICT such as the emergence of WSANs, RFID, M2M data communications, positioning systems and NEDs. These technology developments are considered to be enablers of the vision of an all encompassing infrastructure that bridges "the realms of bits and atoms" (Greenfield, 2006). Many researchers are about to push the evolution of the Internet forward, especially by contributing to the realization of the vision of the RWI dimension of the Future Internet (4WARD, 2010; Beecham Research, 2008; EU-MESH, 2010; Haugli, 2009; ISSNIP, 2010; Pachube, 2010; SENSEI, 2010a; Socrades, 2009; uID Center, 2010; WiSeNts, 2006; and many more).

Most interviewed experts and researchers agree with Haugli (2009) on the assumption that there will be a shift "from closed vertical solution to open horizontal layered value chains".

Among the research activities that design and develop technology platforms, the SENSEI project (SENSEI, 2010a) is outstanding as it is the biggest Integrated Project from Call 1, Challenge 1.1: "The Network of the Future" with an effort of about 1900 person-months and involvement of 19 consortium members from eleven European countries. The SENSEI framework aims at "integrating the physical with the digital world of the network of the future" (SENSEI, 2010a). To make this vision become reality, sensors, actuators, and NEDs have to be deployed everywhere and interconnected. Indeed, WSANs are already deployed on a noticeable scale. However, almost all of them are currently used for single-purposed vertical solutions. Horizontalization is therefore the key design goal of the SENSEI framework, which means the reuse of sensing, actuation and processing services for different applications. The open, business-driven architecture of the SENSEI middleware platform integrates heterogeneous WSANs and provides uniform access to RWI service components. SENSEI services like discovery services or query services enable the development of e-business applications on top of the middleware (Figure 1).
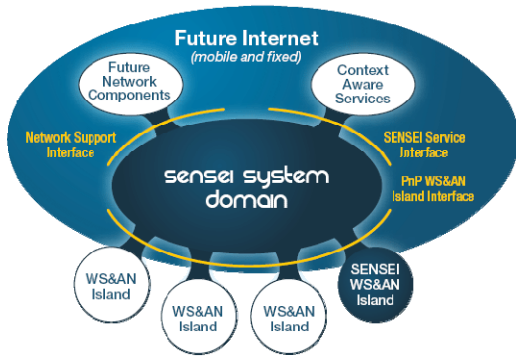

Figure 1: The simplified SENSEI framework.

The SENSEI architecture is designed in a way to satisfy the demands of multiple players in an evolving Internet value network and therefore addresses several aspects, amongst others: privacy and security (protect the privacy of users and offer adequate security for participating systems and entities being observed and acted upon), accessibility (reduced complexity of accessing sensing and actuation services for applications), scalability (support efficient internetworking of a large number of distributed sensor networks islands),

heterogeneity (accommodate a variety of sensor and actuator networks with different technology and administrative domains), manageability (permit distributed management of participating systems and resources), and simplicity of participation (ease the integration of new WSANs) (SENSEI, 2010b).

The horizontal architecture of the middleware platform should enable the reuse of real world services in a uniform manner, which is realized by the means of standardized resource descriptions and flexible service composition. The platform should offer one modular framework with real world services at different levels: on the first level, there are basic sensor and actuator access services, on the second level, basic discovery or sensor and actuator services are offered, the third level, should provide real world context and actuation management services, and on the fourth level, there are dynamic real world service instantiation and composition. The integration platform may have the ability to create a business ecosystem with different players offering any subset of those services. The architecture even allows small providers with low-value transactions to join the framework. This will create new e-business opportunities and contribute to bridging the gap between the net and the real economy. The interviews revealed that the stakeholders expect the concept of the middleware integration platform to enable innovations, support the innovation dynamics within companies, and lead to changes in the market.

4 E-BUSINESS APPLICATIONS

On the basis of the concept of horizontalization, middleware platforms like SENSEI that aim at integrating formerly separated WSAN islands and incorporate M2M data communications have the potential to enable or at least improve a variety of e-business applications. On the basis of such middleware integration platforms, both B2C as well as B2B e-business applications can run, providing business opportunities in different service sectors like retail, public transport, asset management, building automation, smart grid, emergency management, healthcare, or automated manufacturing. Only the application domains asset tracking, building and home, and entertainment can be presented in more detail in this paper.

4.1 Asset Tracking

The need of middleware integration platforms

becomes particularly obvious in supply chains, where goods flow through different organizations, each with its own information system. Current supply chain infrastructures are limited to RFID and barcode technologies and do not yet include sensor data on a large scale. Middleware integration platforms aim at enriching current applications with sensed physical data from various entities and can thus enable e-business applications to reduce delivery delays, improve compliance, secure product delivery to consumers, reduce loss or theft of goods, minimize waste, reduce the extent of management-by-exception, allow ongoing real-time analysis, and ensure supply chain integrity. E-business applications for asset tracking on the basis of middleware platforms are especially motivated by the anticipated achievement of the business benefits: efficiency and quality, and supply chain integrity.

Efficiency and Quality: Sensor data on transportation can enormously improve the efficiency and quality of the underlying business processes as sensors can contribute to full visibility of the supply chain. This allows companies to know exactly where their products are, to analyze ongoing processes, and thus to continually redesign processes and norms. Knowing about products' location and conditions helps to reduce delivery delays as well as to identify and eventually reduce loss and theft. Waste can be minimized, for instance, if the ripest product can be delivered first. For companies this means that fewer products need to be discarded, which leads to lower cost. Moreover, existing infrastructures are limited when it comes to reacting upon deviations between real world data and plan data. Middleware integration platforms that combine WSANs and enable M2M interactions could be designed in a way to immediately trigger actions (via actuators) to counteract deviations between real world data and plan data. Companies can implement proactive processes that are triggered by actuators and thus become a proactive instead of a reacting organization.

Supply Chain Integrity: During the products' flow through the supply chain, companies are faced with compliance requirements, like external regulations imposed by external authorities (e.g., legal, tax, and transport), environmental (e.g., waste electrical and electronic equipment guideline) or social regulations (e.g., business code of conduct). Compliance with these rules in an economical way is essential for companies' competitiveness. The ability to seamlessly trace supply chain integrity protects consumers from rotten food or flawed medicine and gives manufacturers the proof of compliance. Transports sometimes require that

shipments get consolidated into cooled trailers with temperature registration. The information from recent temperature registering devices was only available after the trip. Now, middleware platforms can gather real-time information about current state, location or temperature via sensors that are deployed throughout the supply-chain in order to compare that data with compliance-based planning data and to detect deviations between the 'to be' and the 'as is' state. In case of critical deviation, information systems that are connected to WSANs can initiate appropriate actions to re-establish the supply chain integrity and to prevent further implications of the deviation.

4.2 Building and Home Automation

In the building and home automation application sector an increasing number of objects can be monitored, like water, gas, and electricity consumption or automated, like heating, lights, shutters, or blinds. Innovations such as automatic collection of consumption and status data from metering devices are commonly termed as "Advanced Metering Infrastructure" (AMI). The combination of data from AMI with sensor data, e.g. presence of people, and data from information systems can facilitate various e-business applications. Middleware integration platforms can gather the data that organizations need to integrate from different vertical solutions, share information with other parties, and eventually outsource non-core infrastructure and thus lower cost (see also Haugli, 2009). E-business opportunities are mainly related to applications that help to reduce energy and cost as well as to increase security.

Consumption and Cost Reduction: Recent developments in miniaturization of sensors and actuators and lower prices for bandwidth have increased the potential for integration of different systems: e.g., plausibility checks decrease the number of false alarms, integration of weather forecasts with heating systems reduce energy use. Reduction of energy is especially important in places where people are not too concerned about energy billing as in offices or in public buildings. A probably well-known example of reducing energy is to combine information from a calendar about occupancy of a meeting room with the central heating and lighting system. Temperature and light only need to be convenient in times of meetings in that room. Consumers may benefit from better notification and control of consumption. Concerning water and energy they could realize the concrete consumption of a washing machine and they can be

242

informed about leakage, waste, or malfunction use, e.g. in toilets.

Increased Security: The integration of AMI and WSAN data into one middleware platform can help to detect leakages. Water or gas leakages are difficult to detect because most tubes are underground. The detection of gas leakage is especially important to happen in real-time for security reasons. The integration of data may also help to notice a loss of electricity. Electricity theft occurs in slums of big cities and creates safety risks, e.g. fire or blackouts caused by short cuts.

4.3 Entertainment

In Western societies many people spend a lot of time in entertainment activities such as sports and video games as well as money on equipment and accessories. Consumers seem to appreciate new applications like new games or personalized training programs. A key adoption driver for middleware integration platforms is the fact that several sensors are already deployed for other purposes, e.g. weather forecast, in urban areas. The integration of and ubiquitous access to sensed information of different kind and the connection to information system can pave the way and add value to sophisticated mass customization services. Modern e-business services may satisfy two major customers' needs and wants: real-world assistance and greater amusement.

Real-world Assistance: Integration of sensed information and services like geo-positioning into a middleware platform together with connections to information systems, allows e-business applications to be implemented that users can be supported when moving through unknown cities. Weather forecasts, notifications about traffic jams, or write-ups from other users and recommendations that pop-up in real-time when passing by, might become applicable on cell phones to ease the get along and enrich the experience of visitors. Training plans may be aligned with weather forecasts and modified, e.g. if the air pollution level is too high at the initial planned training session.

Greater Amusement: Video games may be extended to the real world, e.g. scavenger hunts on cell phone based on geo-positioning services.

5 IMPLICATIONS

The deployment of middleware integration platforms and the realizatio of horizontalization may have different implications on business:

Application service providers and WSAN owner can address a broader market: application providers who base their services on top of a middleware platform can compose more sophisticated applications by consolidating information that stem from formerly disconnected networks like WSAN islands (cf. Figure 1). Currently, WSAN are mainly deployed for single purpose use and provide only some niches with data. Through the connection to a middleware platform, single sensing, actuating, and processing services from WSAN islands can be reused for multiple services.

New business roles may emerge: at the moment no "sensor data repository" exists. A new role may be the one of a sensor information broker, some kind of "Google for sensor data", who pre-collects information, caches and re-sells it. The revenue generation mechanism may work by pay-per-use. A similar new role could be a kind of apps store that offers a platform on which new applications can be developed and bought. The business model could imply an open innovation approach, similar to the one of Apple's iPhone apps store.

Rivalry among competitors may increase: Middleware platforms that aim at integration on a global scale provide business opportunities to competitors worldwide. In accordance with Porter (2008) rivalry will increase due to: (1) augmented bargaining power of customers because of higher transparency. Consumers gain ever more transparency about advantages and disadvantages of providers as well as about the quality of services. Internet forums and blogs foster this development, (2) threat of substitute products and services, because higher penetration of mobile services, emerging technologies, and paradigm shifts will provide new product and service possibilities, (3) bargaining power of suppliers, as increased volumes but falling prices may encourage suppliers to cut out intermediaries (disintermediation) or at least strengthen their bargaining power, especially if middleware platforms facilitate direct connections between business partners, and (4) threat of new entrants as the architecture of middleware integration platforms allows small providers with low-value transactions to join the framework.

6 CONCLUSIONS AND OUTLOOK

The sketched implications may challenge the business models of ICT companies as changed assumptions about the environment of an

organization entails changes to the "Theory of the Business" (Drucker, 1994). In the context of middleware integration platforms being deployed and presently isolated networks being integrated, organizations may have to re-define their role and their business scope. It will be a challenge for ICT organizations to assess which business opportunities will generate revenue. This will depend on the attitude of consumers and their willingness to pay for the services, as well as on how middleware platforms will be implemented. Future research is necessary to identify business opportunities that really pay off and to define the companies' roles within a value network in which the concept of horizontalization is realized.

Further challenges and future research topics include: first, the assessment of the extent to which businesses are willing to apply WSAN and M2M technology. Incentive schemes might need to be elaborated to get everyone involved who is required. Second, companies must be empowered to share information via middleware platforms. Risks, which go along with the inter-organizational sharing of data must be identified and minimized. Third, standardization and interoperating among devices is needed. On the one hand, it is important to enable consumers to easily plug in devices and on the other hand it is a prerequisite to reap business benefits due to economy of scales. Forth, security and privacy aspects must be taken into account. The privacy of users must be protected and the adequate security mechanisms must be provided for the participating systems and entities being observed and acted upon. Approaches and solutions to these research topics will foster the evolution of the Internet and extend its business potential.

ACKNOWLEDGEMENTS

REFERENCES

4WARD, 2010. *Architecture and Design for the Future Internet*, http://www.4ward-project.eu.

Amit, R., Zott, C., 2001. Value Creation in E-Business. In *Strategic Management Journal* 22(6/7), 493-520.

Beecham Research, 2008. *Business Opportunities from Remote Device Management: M2M and the Internet of Things*. http://www.beechamresearch.com/infopage.aspx?pageid=15.

Bryman, A., Bell, E., 2007. *Business Research Methods*. Oxford University Press, USA.

Buxmann, P., Hess, T., Ruggaber, R., 2009. Internet of Services. In *Business & Information Systems Engineering* 1(5), 341-342.

uID Center, 2010. *Ubiquitous ID Center*, http://www.uidcenter.org/index-en.html.

Drucker, P., 1994. The Theory of Business. *Harvard Business Review*, September-October, 95-104.

EU-MESH, 2010. *Enhanced, Ubiquitous, and Dependable Broadband Access using MESH Networks*, http://www.eu-mesh.eu.

European Future Internet Portal, 2010. http://www.future-internet.eu.

Floerkemeier, C., Langheinrich, M., Fleisch, E., Mattern, F., Sarma, S., 2008. *The Internet of Things: First International Conference*, IOT 2008, Zurich, Switzerland, March 26-28, 2008, Springer. Berlin, Germany.

Greenfield, A., 2006. *Everyware: The Dawning Age of Ubiquitous Computing*. New Riders. Berkeley, CA.

Haugli, H.-C., 2009. Telenor Objects Vision Paper. http://www.telenorobjects.com/media/1635/telenor%20objects%20vision%20paper%20_190609.pdf.

ISSNIP, 2010. *ARC Research Network on Intelligent Sensors, Sensor Networks and Information Processing*, http://www.issnip.unimelb.edu.au.

Pachube, 2010. *Connecting Environments, Patching the Planet*, http://www.pachube.com.

Porter, M., 2008. The Five Competitive Forces that Shape Strategy. *In Harvard Business Review*, January 2008.

Presser, M., Daras, P., Baker, N., et al., 2008. *Real World Internet: Position Paper*. http://rwi.future-internet.eu/images/c/c3/Real_World_Internet_Position_Paper_vFINAL.pdf.

ReadWriteWeb, 2009. *Top 10 Internet of Things Products of 2009*. http://www.readwriteweb.com/archives/top_10_internet_of_things_products_of_2009.php.

SENSEI, 2010a. *Deliverable 1.3: Business and Social Acceptance Studies for Open and Enterprise Solutions*. http://www.ict-sensei.org.

SENSEI, 2010b. *Integrating the Physical with the Digital World of the Network of the Future*. http://www.ict-sensei.org.

Socrades, 2009. *Service-Oriented Cross-Layer Infrastructure for Distributed Smart Embedded Devices*, http://www.socrades.eu.

Timmers, P., 1998. Business Models for Electronic Markets. In *Electronic Markets* 8(2), 3 - 8.

Venkatraman, N., 1994. IT-Enabled Business Transformation: From Automation to Business Scope Redefinition. In *Sloan Management Review* 35(2), 73-87.

WiSeNts, E. (2006). *Cooperating Embedded Systems for Exploration and Control featuring Wireless Sensor Networks*. http://www.embedded-wisents.org.

SHARE VS. OWN
Software Reuse using Product Platforms

Paul Abraham, Vishal Sikka

Office of the CTO, SAP AG., Dietmar Hopp Allee, NeurottStrasse16, D-69190 Walldorf, Germany
Paul.Abraham @sap.com,Vishal.Sikka@sap.com

Gordon Simpson

Independent Enterprise Technologist, San Francisco, CA, U.S.A.
g.simpson@sap.com

Abstract: SAP is a complex multi-national development organization with a large number of diverse products and changing target markets. Effective allocation of resources is a difficult at the best of times. Of late, the target markets, and supporting technologies, change every couple of years exponentially increasing the complexity, necessitating a way of recalibrating that keeps pace with new realties. SAP, with a mature understanding of functional, software and technical relationships, has adopted a platform approach covering both functional and technology capabilities. However, a variety of factors, many in the management space, prevent that from being effective. This paper will explain why product-line/platform is a better strategy than platform or custom product strategies, in a way that can be understood, proven and adopted by management and developers alike. Specific recommendations of practices for delivering reuse effectively are also provided.

1 PROBLEM CONTEXT

SAP spends a great deal of time examining how to allocate resources among its development groups. This allocation is made most efficient by sharing resources among products; dedicating resources only when dictated by necessity. The advent of groundbreaking technology changes or changes in the marketplace force us to re-evaluate this allocation. The acceleration of these previously slow moving factors requires SAP to develop a systematic way of matching pace with evaluation. These change factors and associated resource issues are, to some degree, applicable in turn to SAP's ecosystem i.e. development partners, technology vendors and customers.

When, how and what to share, of software components, processes, infrastructure, people and knowledge, will need to be continuously evaluated as SAP products go through several cycles of renovation. SAP has thus far adopted reuse primarily through a product platform strategy, using both technology and applications platforms in its

products. This paper presents the business reasoning for reuse of software and processes in a product-line/platform vs. dedicated to a product and attempts to fully understand the reasoning and economics behind making such a decision.

Practice has shown that many inhibitors lie in the realm of management and decision-making. The most common causes of these failures are: resource constraints, lack of incentive, single-project view, time constraints, lack of clarity on reuse utility, and lack of education. Software project managers and developers need to achieve better understanding, estimation, evaluation, and quantification of the software reuse and associated business factors as well as their predictive relationship to software effort and quality.

Contemporary software reference models for reuse do not consider many of the technical and non-technical factors in their quantitative models. To mitigate this, the paper also explores the broader, industrial engineering perspective and its concepts of product platforms, product lines and other relevant methodologies to proposes strategies for

designing and building reusable application software components.

We present a rigorous, cost benefit analysis based methodology for the evaluation of well-defined metrics to measure the benefits of a particular strategy and the associated costs. We recognize that the decision-making is essentially probabilistic, leveraging imperfect data, making such decision-making shades of gray among alternatives rather than black and white. The presented methodology accounts for these imperfections in the analysis.

Guidelines are developed to help a decision-maker decide when the long-term benefits involved in implementing and maintaining reusable coding procedures outweigh the short-term benefits of a dedicated implementation. Specific recommend-ations are made for coding practices, software design, documentation and management procedures that encourage and result in successful code reuse practices.

2 SOFTWARE REUSE INVESTMENT SUCCESS FACTORS

The industrial (manufacturing) world has been successful for many years in implementing a product-line approach to reuse using pre-fabricated (pre-manufactured, interchangeable) components. Product-line is closely related to the concepts of horizontal and vertical reuse. Horizontal reuse provides generic reusable components that can support a variety of products. Vertical reuse focuses on developing the preferred parts supporting a given family of related products or product-line. It is regular practice for these industries to assemble parts into products and use the same parts in more than one product within a "product-line" family.

Can this same "manufacturing" approach be used in software engineering?

SAP has certainly embraced parts of this philosophy with its Business Process, Application and Technology Platform strategies that serve as the foundation on which our *Business Suite* (Suite) and *Business By Design* (BYD) products are built.

2.1 Managing Diverse Software Products

To run any software component requires the use of other software and hardware artefacts that may be owned by SAP or shared among its partner ecosystem. *It is reiterated that, for this paper, the topic of reuse also includes artefacts used in the run time for the software components.* Assuming we have the relevant tools, processes and technologies for effective software component reuse, the evaluation of whether to reuse a software component (share) or develop it (own) is dependent on the particular reuse strategy we adopt.

There are two dimensions to the strategy.

- On one hand, we look at the product platform strategy and decide where a particular component should reside: in which architectural layer, using what technology etc.

- On the other hand, we look at the time horizon for the business initiative the software component is in support of: i.e. today's business, the next generation of emerging businesses, and the longer-term options out of which the next generation of businesses will arise.

At SAP, we have adopted multiple strategies depending on the nature, size, location and technology associated with the software component being reused. We briefly outline the problem environment in the graphic below:
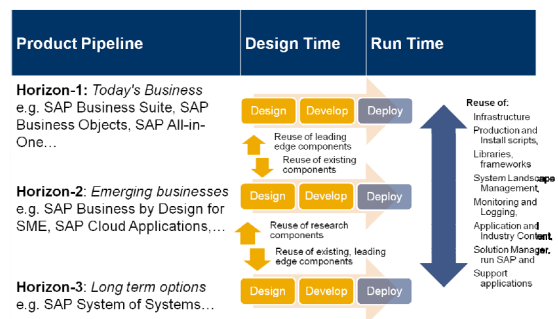
Figure 1: Shared vs. Own problem environment.

Note that there are two cycles in play. Across the functional cycle of design, develop, deploy, components change based on fit to purpose with upgrade and rebuild as the scope of the purpose changes over time. The vertical cycle is a much longer and slower moving one based on use density. As components are more frequently reused, they sediment down through the platform layers. This sedimentation can often include 3rd party infrastructure layers, who adopt technical capabilities initially developed for a single application solution. In the most effective model the

reverse flow is also managed, with components being factored out of the platform and back into products as the reuse decreases over time. This allows for the complexity and overhead of the platforms to remain optimal over time.

2.2   Balancing Commonality and Distinctiveness

At a fundamental level, product variety and fit is valuable in the marketplace. The need for superior performance of the products and the desire to preserve distinctiveness (e.g. custom features, control etc.) promotes product organizations to own certain key components. On the other hand, it is costly to deliver, as cost benefits are driven through commonality. The balanced sharing of assets across products allows companies to manage this trade-off.

This balance has however, temporarily resulted in an unwanted side effect in SAP: total cost of ownership (TCO) increases due to the complex configuration that we provide customers to tailor our software products to their needs. As mentioned above, parameterization is a valuable tool in leveraging shared assets to fit different solutions. However, the current architecture leverages the same parameterization for both SAP engineering product fit and on-premise customer fit (customization). The effect is to trade off customer complexity for the power of reusing, and therefore only having to support, a single mechanism.

Changing the product architecture can influence the nature of the trade-off. For example by the use of pre-configured and interchangeable software component for a particular industry vertical or customer group which hide the complexity of configuration will lower costs of customization. Another technology solution is the use of model-based methodologies that lower the fixed cost of developing software, and/or delivering the software as a service. The hypothesis is that this type of reuse promotes mass-customization, shortens the time to market and promotes consistency in products.

2.3   Common Architecture Strategy

The sharing or owning of software components is dependent on the architectural strategy. The architecture relates software components to a physical problem space (hardware, operating system, and application packages such as database or user interface). A common architecture lessens the need to make reusable software components highly generic because the environment in which they will be used is well defined. The architecture defines the rules for developing software components and provides standard interfaces and data formats. This aids in the inter-changeability of reusable software components across the product-line.

SAP had elements of a common architectural strategy from its inception. SAP uses this approach for the lower level technological platform and to support the user interface. However, there is considerable difference in higher layers of the architecture between our Business Suite and Business by Design (BYD) products, which leverage the same technology platform but are targeted at different markets. It should be reiterated that the platforms should be different if they are fundamentally different and that the determination of that "fundamental difference" is at the heart of SAP's challenges.

2.4   Product Platform Strategy

Product platform strategy is the foundation of the existing SAP product strategy, which has multiple products related by common technology platform. It defines the cost structure, capabilities, and differentiation of the resulting products. When the market and products were less diverse, and the technology considerations more unified, separating product platform strategy from product line and individual product strategy allowed SAP to concentrate on its most important strategic issues of reliability and scale. As the diversity and rate of change has increased, the question as to which components products share and which are dedicated has ultimately tied to the product platform strategy more closely to the product line.

2.5   What is a Software Product Line?

A *software product line* is a set of software-intensive systems, satisfying the specific needs of a particular market segment, that share a common, managed set of capabilities and that are developed using a common methodology and leveraging common skills sets.

This definition is consistent with the traditional product line definition. But it adds more: it puts constraints on the way in which the systems in a software product line are developed. Substantial production economies are achieved when the systems in a software product line are consistently developed from a common set of assets in contrast to being developed separately, from scratch, or in an arbitrary fashion. It is exactly these production

economies that make the software product line approach attractive.

Production is made more economical when each product is primarily formed from existing components, tailored as necessary through pre-planned variation mechanisms such as parameterization or inheritance, adding any new components only when necessary, and assembling the collection according to the rules of a common, product-line-wide architecture. Building a new product (system) becomes more a matter of assembly than one of creation; the predominant activity is integration rather than programming. For each software product line, there is a predefined guide or plan that specifies the exact product-building approach.

Software product lines give *economies of scope*, which means that we take economic advantage of the fact that many of our products are very similar–not by accident, but because we planned it that way. We make deliberate, strategic decisions and are systematic in effecting those decisions. This concept must be contrasted with the specifics of a product platform, which is described next.

3 THE PRODUCT PLATFORM

We define a product platform as a collection of core assets that are shared by a set of products. These assets can be divided into four categories:

Software components – A software component is a unit of composition with contractually specified interfaces and explicit context dependencies. A software component can be deployed independently and is subject to composition by third parties.

Processes and infrastructure - used to make or to assemble software components into products

Knowledge base – design know-how, mathematical models, testing methods and data sets

People and relationships – teams, relationships, between members and between teams

In certain manufacturing systems, these process and systems are themselves machinery like assembly lines or manufacturing centres etc. In the software arena new ideas like *software factory* embody this principle. Most companies do have parts of this automated with production and installation scripts, configuration of system landscapes etc. However, a coherent methodology and infrastructure is yet to emerge. The organizational aspects also need to be

facilitated by automated systems much like manufacturing centres. This encompasses the knowledge base since a large part of the knowledge resides in people. Web 2.0 holds a lot of promise in this area and is being integrated into development and production tooling.

A product platform is primarily a definition for planning, decision-making, and strategic thinking. A product platform is not a product; it is a collection of the common elements, especially the underlying defining technology, implemented across a range of products. So in a sense this definition is broad, a generalization of the concept in SAP where we have of a technology platform and application platform but in another sense it is distinct, as it results in a collection of common elements. These common elements need not necessarily be complete in the sense that they are something that could be sold to a customer.

SAP markets and builds its products (Suite and BYD) as platforms for running the business processes of large enterprises (LE) and small to medium enterprises (SME) respectively. The defining technologies used to implement this business process platform will evolve over time, at different velocities and hence it is imperative to manage this effectively. The platform's unique differentiation provides a sustainable competitive advantage. Therefore it may be argued that all components that are related to business processes such as orchestration must necessarily be part of the platform and cannot be owned by an application or industry solution built on top of it.

3.1 The Platform Influence on SAP's Ecosystem

In the context of SAP's ecosystem, a platform may be viewed as a realization of the technology strategy that is made available through a set of access points or interfaces (APIs). Partner ecosystem members (ISV's and SI's) then leverage these interfaces as a kind of toolkit for building their own products and solutions, and think of them as the starting point for their own value creation. The platform is the "mechanism" through which the platform organizations share value with their ecosystem. Any product contains elements specific to a given use or solution and elements that are shared with many other products in the development ecosystem within and outside SAP development. The latter represent an opportunity that can be leveraged by other members of the ecosystem to eliminate redundant effort.

The architecture of products and services has a profound effect on the evolution of ecosystems. Well-managed platforms shape ecosystem dynamics as they grow to incorporate new functionality and create opportunities for SAP to expand its ecosystem. How a platform evolves and responds over time, shapes the ecosystem that depends on it: what firms survive, where diversity can exist, what will be easy to do and what will be hard, which things in the ecosystem will do well with little effort, and which things will be challenging. This happens because platforms serve as an intermediary between the underlying technology and the ways in which it can be easily exploited.

3.2 Platform Architectures

The choice of a defining technology as a platform strategy is perhaps the most critical strategic decision for a high-technology company. Typically, the defining technology of a platform differentiates the products that are based on that platform. While SAP is more defined by its business process centricity, the technology aspect of the platform strategy is still significant. Business applications can be classified in different operational or technical archetypes, based on their characteristics and requirements at run-time. A few examples of these archetypes are:

- **Online transaction processing systems (OLTP)**: characterized by low latency, high responsiveness, data integrity, predefined UI workflows. Instances of this archetype are e-commerce sites, CRM, e-banking systems.

- **Analysis systems or online analytic processing (OLAP):** characterized by their ability to produce complex analytical and highly customizable queries on large multidimensional datasets, with low latency responses. Business Intelligence (BI) systems fall into this category.

- **Batch systems**: capable of performing operations on large datasets efficiently, coordinating jobs to maximize CPU utilization and energy consumption with recovery policies when exceptions occur.

- **Networked systems**: software that integrates different applications and services into more complex solutions. It differentiates itself by delivering a business solution (e.g. Supply Chain Management) that manages the information and control flow across many other

systems (Inventory Management, Order Processing).

SAP has products or components that cover all of these archetypes. Each of these application families has its own constraints, characteristics, and optimal design patterns that can be applied to solve the specific challenges they present. Very often, these challenges have conflicting goals. For example: OLTP will optimize for low latency, whereas latency for batch or networked systems is not as important. OLTP scales better horizontally and benefits from a stateless architecture, while batch systems scale vertically and tend to be stateful. The technical infrastructure and services to support each is consequently significantly different. The key point is that a platform's effectiveness is highly dependent on the archetype served. The more knowledge of the application a platform has, the greater its ability to increase the efficiency of running and operating it, and the greater the degree of sharing. Thus having multiple platforms that are tailor made to its constituency (by archetype, mode of delivery or size: LE, SME Micro) one could substantially lower the TCO.
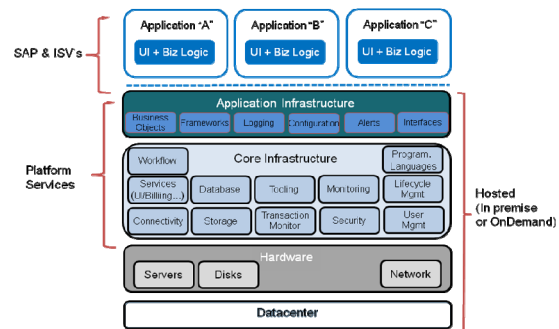
Figure 2: Increased reuse through an application runtime infrastructure.

An increased amount of shared components leads to higher levels of efficiency, so the question is which are the most natural candidates to be "extracted" from applications into the platform? The obvious candidates are those referred to application infrastructure services *viz*. application configuration, run-time exception, logging etc. Refer example above. Every application needs them, yet they are frequently written repeatedly for each platform.

By exposing these basic services publicly or by sharing them across platforms as libraries or frameworks, the platform has an increased ability to automate common procedures and offer more advanced operational management capabilities and lower the TCD by sharing the common platform

services. Thus, finer-grain tuning, customization and troubleshooting are made available. In the example above, notice that the hosting method does not need to understand in detail what the application does, but instead how it does it. (*e.g.*, where are connection strings to the database stored? How is run-time exceptions logged and notified?)

4 COST-BENEFIT ANALYSIS FRAMEWORK

Finally, given one or more credible product platform strategies for reuse, we would like to evaluate these alternatives to estimate the immediate economic benefits of reuse. The following CBA methodology is adapted from various publications and is presented for illustrating the benefits of a platform strategy and for completeness.

4.1 Metrics to Measure Benefits of Reuse

In this section, we present a conventional cost benefit analysis. Benefits of reuse are difficult to measure objectively and hence we present the current state of the art for it. Cost on the other hand is easier to compute and only a high-level view of what is pertinent pointed out. In addition, enterprises have developed their own methodology for cost accounting, which are uniform in spirit but not in detail. Software metric is any measurement that relates to a software system, process, or related documentation. Metrics are distinguishing traits, characteristics, or attributes that are both static and dynamic. The reuse metrics mentioned below are those developed by research teams at George Mason University in Fairfax, Virginia (see Rine & Nada 1998). These relate to the benefits of software reuse in the cost benefit analysis calculations that is introduced further down the paper. We also use the term module in place of software component as the basic unit for applying metrics. The distinction is somewhat academic but a software component may be too coarse grained for meaningful measurements.

In order for a metric value to be statistically valid, it is necessary to have a reasonable quantity of data. This data collection is unlikely to be successful unless it is automated and integrated into the development process. Finally, product data should be kept as an organizational asset and historical records of all projects should be maintained. Once an appropriate data set is available, model evaluation involves identifying the parameters that are to be

included in the model and calibrating these using existing data. Such model development, if it is to be trusted, requires significant experience in statistical techniques.

The software reuse metrics are grouped into five major categories: *general, quality, parameterization, coupling, and cohesion,* which are also software engineering principles that correspond to the reuse attributes. The general category is for attributes that are not in the four software engineering categories. The following table lists the popular reuse metrics:

Table 1: Metrics to Mesure benefits of Reuse.

Category	Metric	Measures
General Metrics	Time to market	Reduction in development time
	New Product Opportunity Opportunities	Potential for long tail solutions etc.
	Understandability	Ease of adoption
	Size	Extent of the module
	Type of module	e.g. Specification, functional
Quality Metrics	Consistency	for centralized maintenance
	Ease of Change	The degree to which it can be changed
	Comments	Usefulness, understandable. accuracy
	Formatting	Readability of the code
Parameter-ization	Functional Data	
Coupling Metrics	Coupling System External Coupling	the strength of the interconnection and dependency among modules
Cohesion Metrics	Functional cohesion	degree to which each part of the module is necessary for performing a single function.
	Data cohesion	degree to which it has a single-data type associated with it

4.2 Cost Calculations

The methodology proposed here is an amalgamation of methodology proposed by the Software Institute in HP Labs and the Software Engineering Institute in CMU (see Clements et al., 2005, Petersen, 2004 and Malan & Wentzel, 1993).

4.2.1 Development Costs

Setup and Overhead. New systems will be needed to support a full-fledged systematic reuse program. It includes ongoing costs of expanding and maintaining reuse layer or system, a management support structure to ensure systematic reuse, and training programs, and should be assessed as indirect overhead.

Producer. The reusability of software components depends on a number of factors such as the degree of generality, complexity, and fit to expected use, as well as the quality of the component, and the extent and utility of documentation and accompanying test suites. Further, the component has to be available, and hence must be certified and entered in a platform, library, as software services or broadcast by some other means. Therefore, component producers face additional costs over and above the usual development-cycle costs, and these are estimated to be anywhere from 30% to 200% higher than the cost of producing a component not intended for reuse. This is true even for a component that is re-engineered from existing code.

Consumer. Selection, specialization and integration for reuse entails articulation of the component requirements in a suitable form, search and retrieval of the component, understanding of what the component does, and verification that it does indeed fit the purpose. The component may need to be specialized to fit the consumer's current needs. This involves adaptation (with co-requisite program understanding and subsequent testing). Lastly, the component must be integrated into the system under development, and tested.

Lifecycle Costs. The view of maintainability as a form of reusability is novel and important. It captures the idea of reusability in time within a dynamically evolving system. Evolutionary dynamic systems require reusability in time of unchanging parts of the system while other parts of the system evolve. By centrally maintaining the reuse components, managing their evolution, and propagating upgrades to new products as well as updated versions of older products, the organization can exploit further opportunities to reduce duplication of effort. Moreover, centralized enhancements to black box components enable a whole platform of derivative products to be produced more quickly at lower cost.

4.2.2 Probabilistic Nature of Calculating Cost

Time value of Money. When the reuse instances are expected to occur over a longer time horizon, the timing of the cash flows should be taken into account. This is done by incorporating a standard present value analysis into the model. One typically uses the Discount Cost Function (DCF) analysis. Typically Horizon-1 (today's businesses) components described previously fall into this category. In case of Horizon-2 & 3 (emerging businesses, and the longer term options) components, more market based approach such as real options valuation (ROV) may be used to account for the high uncertainty.

Uncertainty in Reuse Instances. The degree of uncertainty about the evolution of a product family tends to increase as the time horizon is stretched. Thus, anticipated reuse opportunities arising from products or upgrades planned in a multiyear horizon is likely to be much more uncertain than those in the current one-year business plan. To incorporate the uncertainty as to whether the component will indeed be reused, the probability of each reuse instance should be estimated, and the expected consumer savings computed. This is essentially a DCF calculation of the NPV and optionally decision tree analysis .

Future upgrades of Components. The maintenance and management of evolving components increases the cost to the producer/maintenance group. Consumers of the component benefit from not having to duplicate corrective and evolutionary maintenance activities, though they do have to incur some cost to incorporate upgraded component(s) into their products. This also is essentially a DCF calculation of the net present value (NPV).

4.3 Cost Benefit Analysis

It is clear from the discussion above that for some components the choice of strategy dictates whether it is shared or dedicated for a particular product. In many cases however software components may not have a predestined position in the architecture or is not obvious and hence a cost benefit analysis of shared versus own using historical data on similar projects should be performed. Since reuse involves multiple products evolving through their respective life-cycles, an assessment of the economic impact of a systematic reuse program must incorporate cost and revenue projections that extend beyond that of a single development project. A template such a cost

benefit analysis that should be performed for each component based on the research from the Software Engineering Institute (SEI) at CMU for is presented below.

Table 2: Cost functions used to compare building a reuse platform versus building stovepipe products.

Function	Output
$C_{org}()$	Cost to setup and run an organization to adopt the product line approach for its products
$C_{cab}()$	Cost to develop a core asset base suited to satisfy a particular scope
$C_{unique}()$	Cost to develop the unique parts (both software and non-software) of a product that are not based on core assets
$C_{reuse}()$	Costs to build a product reusing core assets from a core asset base
$C_{prod}()$	Cost of building a product in a stand-alone fashion. It relies on historical data or general software engineering cost models for its evaluation.

Note: We assume that these functions accommodate influencing factors such as the time value of money, uncertainty in reuse instances and the probability future upgrades. Specific formula's are available in the research from several institutions like SEI, HP Labs etc. (see citations).

This cost can be expressed by Equation 1.

Cost of building a product line =

$$C_{org}() + C_{cab}() + \sum_{i=1}^{n}(C_{unique}(product_i) + C_{reuse}(product_i)) \quad (1)$$

This equation says that the cost of fielding a product line is the cost of organizational adoption plus the cost of building the core asset base plus the cost of building each of the *n* products. The cost of building a product is the cost of building the unique part of that product plus the cost of incorporating the core assets into the product. The cost of building *n* products independently, is expressed in Equation 2.

Cost of building *n* stovepipe products =

$$\sum_{i=1}^{n}(C_{prod}(product_i)) \quad (2)$$

4.3.1 Evolution and Upgrade

To account for a cycle of product evolution—that is, the time in which a product appears in a new version, probably with new or at least improved features—under the non-product-line, the model introduces a new cost function, $C_{evo}()$. This function is parameterized with product and version numbers and returns the cost of producing that version. One

might make a first approximation by assuming that the cost to produce a new version is some percentage of producing the original product; for example $C_{evo}()$ = 20% * $C_{prod}()$

To calculate the analogous cost under a product line regime, we introduce a new function, $C_{cabu}()$. This function returns a measure of how much it costs to update the core asset base as a result of releasing a new version of a product. Changes to the core asset base can occur because the new version required changes to or exposed bugs in existing core assets. Changes can also occur when new features expose new commonalities with other products that were considered unique but now can be refactored into commonalities.

4.3.2 Benefits Calculation

Software product lines bestow benefits to the developing organization besides direct cost savings. For example, they often allow an organization to bring a product to market much more quickly. We can accommodate these other factors by using benefit functions that are similar to the cost functions introduced in the basic model. Unlike the cost functions, there is no fixed number of benefit functions. However, the metrics discussed previously help one establish a list of benefits (= nbrBenefits) to be factored in the analysis as given in Equation 3.

Benefits of building *n* products using a product platform approach =

$$\sum_{j=1}^{nbrBenefits}(B_{ben_j}(t)) \quad (3)$$

where ben_j is a specific benefit and $B_{ben_j}()$ is the benefit function for that benefit. Each benefit function is parameterized by the time period of interest since the benefits may vary over time.

The contributions of the benefits are summed and then used to build a model equation as needed. For example, to express the development cost savings (or loss) from using the product platform approach as opposed to one-off development for each product equals [Equation 2] – [Equation 1]. A more complete picture of the cost benefit of using a product platform approach adds Equation 3 to that result.

4.3.3 Illustrative Examples of Reuse in the Design Time

The following example illustrates the impact of the approach in a new product line. The new product

line is targeted for the SME space and consists of about five components called distribution units (DU). Since this product line was targeted for a business user and in an effort to maintain the look and feel across product lines, it was decided that the current technology platform use for LE space called NetWeaver is the appropriate platform for this product line. However, experience has shown that there was potential to reuse a lot of features and services among the five DU's. Thus it was decided another (sixth) component called the application platform was to be developed by a producer group for this product line.

Challenge. Beyond building the initial business case, there is normally a low confidence level in any data related to the future success and adoption timeframe of a new software product line. This tends to drive the design decision making away from reusable/platform towards single use components - "get the first product out of the door and worry later". This can have potentially damaging effects on the actual ability to grow the product line in a cost effective manner but has always required subjective judgement on the part of the solution manager. Use of the CBA would allow the initial solution managers to "run the numbers" associated with making certain "reasonable estimates" around different componentization strategies and the short, medium and long term financial impact.

Solution. Based on measured data obtained from many prior software development projects, estimates show that with a 50% reuse level and a 5x quality improvement in the reused component over new code. However, Producer effort was increased by 108% and consumer effort reduced by 40% during the development phase. During the maintenance phase it is estimated that producer effort was increased by 25%. and consumer effort reduced by 42%.

These effort factors (assuming a 50% reuse level) are used together with the following assumptions to estimate reuse benefits:

- Hourly rate for software engineers (including basic salary $75 and administration overhead)
- Project team size 20
- Development cycle time without reuse (months) 12
- Annual inflation in labour rate 5%

A simplifying assumption that all of the products are comprised of the same amount of new and reused code is made to better demonstrate, a number of points. We have also simplified the cost benefit calculations and used only NPV calculations. We divide the calculations into four models that have self-explanatory titles. The model results are shown in the table below.

Table 3: Examples of Reuse in the design time.

Model 1: Basic Development Phase Costs

	Producer Cost	Product1	Product2	Product3	Product4	Product5
Year of Release	0	1	1	2	2	3
Without Reuse	300,000	630,000	630,000	661,500	661,500	694,575
With Reuse	624,000	378,000	378,000	396,900	396,900	416,745
Reuse specific Overhead		35,000		25,000		25,000
Consumer Saving		252,000	252,000	264,600	264,600	277,830
Cumulative Net Saving	(-)624,000	(-)407,000	(-)155,000	84,600	349,200	602,030

Model 2: Taking the Time Value of Money Into Account

	Producer Cost	Product1	Product2	Product3	Product4	Product5
Consumer Saving after Discounting [i]		234,419	234,419	228,967	228,967	223,642
Cumulative Discounted Net Saving	(-)624,000	(-)422,140	(-)187,721	19,613	248,580	452,098

Model 3: Taking Uncertainty In Reuse Instances Into Account

	Producer Cost	Product1	Product2	Product3	Product4	Product5
Probability of Reuse		1	1	0.90	0.75	0.50
Consumer Saving with Discounting & Uncertainty		234,419	234,419	206,070	171,725	111,821
Cumulative Discounted Expected Net Saving	624,000	422,140	187,721	3,284	168,441	260,138

---

[i] The interest rate may be the prevailing bank rate, reflecting the interest that the investment would earn if it was deposited instead of invested in reuse, or the company's hurdle rate, reflecting what the investment would earn in some alternative use within the company.

Table 4: Examples of Reuse in the design time. (cont.)

Model 4: Including a Future Upgrade

	Producer Cost	Product1	Product2	Product3	Product4	Product5
Year of Release	1	2	2	3	3	4
Probability of Reuse		1	1	0.68[ii]	0.38	0.25
Upgrade without Reuse	75,000	157,500	157,500	165,375	165,375	173,644
Upgrade with Reuse	93,750	91,350	91,350	95,918	95,918	100,713
Reuse specific Overhead		10,000		10,000		10,000
Additional Consumer Saving		57242	57242			
				38019	21246	13653
Cumulative Discounted Expected Net Saving	(-)717,750	(-)467,301	–175,641[iii]	38,766	231,737	329,599

5 CONCLUDING REMARKS

SAP, the world's largest provider of enterprise applications software, originally architected its reuse strategy around horizontal application and technology platforms that provided focus on the scaling and reliability desired by its mainly homogenous enterprise market. On top of this strategy, it built a large and geographically distributed organization and a large portfolio of diverse products.

Recently, an increased rate of change in market needs and supporting technology innovations has stressed that strategy. Solution managers, development decision makers, are challenged to effectively handle the conflicts of rapid solution delivery while identifying candidate components for application or infrastructure reuse. This complexity extends out beyond the company into its ecosystem of partners and customers as they fit the applications to specific business needs. Thus, guidelines should be developed and specific recommendations made to streamline this evaluation process.

It is our position that a product-line approach, supported by provable cost benefit analysis, is a more effective model for delivering reuse benefits in this dynamic market environment. As software industry models of reuse are not sufficiently robust, we have looked to traditional manufacturing industries for guidance; moulding their models to fit the imperfect data base of software decision making.

This paper proposes a cost benefit analysis based model which, when combined with a methodology, software engineering tooling and organizational guidelines, will enable the engineering management to effectively balance product specific and platform reuse requirements, in a cost and market effective manner.

As this model has not yet been adopted, the paper also describes the steps necessary to fit the proposed approach to a specific organization and how the calculus would provide objective componentization and reuse data. Solution management in the design of new product lines would leverage this. Comparative examples are given covering the first product in a new line, new market segment and the first product in a product line largely similar to an existing line. These two scenarios have significantly different subjective influences, requiring different use of the cost/benefit analysis.

[ii] Conditional on Product 3 being produced, the probability of its upgrade being produced is assessed as 0.75. Thus, the probability of producing the upgrade is 0.9*0.75 = 0.68. Similarly, the conditional probability of upgrades to products 4 and 5 being produced is 0.5.

[iii] This is the cumulative saving for the first product plus its upgrade. Overhead for the year is assumed to be incurred regardless of whether the next reuse instance is actualized (i.e. a semi-fixed cost, such as salaries).

REFERENCES

Clements, Paul C., McGregor, John D., Cohen, Sholom G., 2005. *The Structured Intuitive Model for Product Line Economics (SIMPLE)*, Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA.

Dale R. Peterson, F. van der Linden (Ed.), 2004. *Economics of Software Product Lines,* PFE 2003, LNCS 3014, pp. 381–402, 2004. Springer-Verlag Berlin Heidelberg.

Iansiti, Marco and Levien, Roy., 2004. *The Keystone Advantage: What the New Dynamics of Business Ecosystems Mean for Strategy, Innovation, and Sustainability*, Harvard Business Press.

Malan, Ruth, Wentzel, Kevin, 1993. *Economics of Software Reuse Revisited*, Software Technology Laboratory.

McGrath, Michael E. 2001. *Product Strategy for High-Technology Companies: Accelerating Your Business to Web Speed*, McGraw-Hill.

Rine, David C., Nada, Nader 1998. *Software Reuse Manufacturing Reference Model: Development and Validation*, School of Information Technology and Engineering, George Mason University, Fairfax, Virginia.

Robertson, David, Ulrich, 1998. Karl *Planning for Product Platforms*, Sloan Management Review.

Viguerie, Patrick, Smit, Sven and Baghai, Mehrdad, 2008. *The Granularity of Growth: How To Identify The Sources Of Growth And Drive Enduring Company Performance*, John Wiley & Sons (US).

THE COMETA E-INFRASTRUCTURE
A Platform for Business Applications in Sicily

Marcello Iacono Manno, Pietro Di Primo, Gianluca Passaro, Emanuele Leggio

COMETA Consortium, Via Santa Sofia 64, Catania, Italy
{marcello.iacono, pietro.diprimo, gianluca.passaro}@ct.infn.it, emanuele.leggio@gmail.com

Roberto Barbera

Physics & Astrophysics Department, Catania University, Via Santa Sofia 64, Catania, Italy
roberto.barbera@ct.infn.it

Giuseppe Andronico, Riccardo Bruno, Emidio Giorgio, Marco Fargetta, Giuseppe La Rocca
Salvo Monforte, Diego Scardaci, Fabio Scibilia

National Institute for Nuclear Physics, Catania Section, Via Santa Sofia 64, Catania, Italy
{giuseppe.andronico, riccardo.bruno, pietro.diprimo, emidio.giorgio, giuseppe.larocca, marco.fargetta,
salvatore.monforte, diego.scardaci, fabio.scibilia}@ct.infn.it

Abstract: The COMETA e-Infrastructure running in Sicily is compliant with the EGEE middleware and specifications offering a great computing power and huge storage capacity. Since its beginning one of the main goals has been to extend the adoption of Grid paradigm from the academic to business world. Several software and hardware extensions have been implemented in order to enhance the infrastructure performances; they include a new low-latency net layer reserved to heavy parallel applications; some modifications to the parallel job submission and execution procedures for a better support of MPI-based applications; new tools for job monitoring and file catalogue interaction; a scheduling policy tailored on the requirements of a complex environment hosting heterogeneous jobs; the GridLM license server able to grant access to commercial software only to authorised users; finally, the Secure Storage Service defends from insider abuse and completes a very high security level environment. The proposed business model includes a wide range of services collectively defined as Infrastructure-as-a-Service. All the above components result in a powerful and flexible platform, easy to use for any applications and open to further developments.

1 INTRODUCTION

The COMETA (www.consorzio-cometa.it) Grid infrastructure is a distributed computing platform located in Sicily both for academic and business computing. It is the first in Italy scaled to involve a whole region. Since its preliminary design, a primary target has been the implementation of business applications. For this reason, the environment of Grid computing has been adapted to the demanding requirements of such applications.
Grids can be useful as cheap host platforms for several kinds of business applications ranging from High Performance Computing (HPC) to Cloud Computing (CC), from data intensive to applications requiring dedicated user-friendly interfaces. Particularly, Small and Medium Enterprise can reduce both their cost of ownership and time to market by outsourcing their design and/or production activities. Even Public Institutions (such as hospitals and/or cultural institutions) may take advantage from this solution.

The business model is Software-as-a-Service (SaaS) that evolves into an Infrastructure-as-a-Service (IaaS) as the whole infrastructure is transparently used as a facility providing on-demand services for heterogeneous applications.

The effort to really and effectively support such a scenario required deep adjustments both to hardware and software compared to the standard EGEE project (www.eu-egee.org/) gLite middleware

(www.glite.org) that only provides generic tools to access the bare resources. The following sections describe the efforts about hardware design, policy tailoring and software support focusing on each specific issue. Section 2 introduces the COMETA e-Infrastructure. Section 3 describes the adoption of the InfiniBand net layer. Section 4 is about the scheduling policy, a critical aspect of the management of such infrastructures. Section 5 describes the modifications to the standard official middleware for a better support of MPI-based parallel applications. Section 6 focuses on GridLM, the newly-developed license server allowing only entitled users to run licensed software on the Grid under the license terms and the accounting system. Section 7 describes the Secure Storage Service implemented to solve the insider abuse problem. Section 8 is about the so-called "watchdog" tool to monitor long lasting jobs and recursive commands for massive interaction with the data catalogue. Section 9 illustrates the most common porting procedures and related computing schemas. Finally, some conclusions are drawn in Section 10.

2 THE COMETA CONSORTIUM AND THE SICILIAN E-INFRASTRUCTURE

The COMETA Consortium gathers the Universities of Catania, Messina Palermo , the National Research Institutes for Nuclear Physics (INFN), Astro-Physics (INAF), Geo-Physics and Volcanogy (INGV), and the SCIRE Consortium. These institutions have been involved at different levels in the EGEE project and other European projects deploying international e-Infrastructures. Nevertheless, the idea behind the development of a new regional e-Infrastructure is to foster the adoption of Grid computing for massive computation among Sicilian researchers both from academia and business worlds. For this reason the Sicilian e-Infrastructure has been built on grants coming form the Italian Ministry of University and Research (PI2S2 Project, www.pi2s2.it/) and the Sicilian Regional Government (TriGrid Project, www.trigrid.it). The infrastructure, which adopts the gLite middleware, is fully compliant with the international standards on interoperability. Figure 1 shows the location of the seven sites of the infrastructure in the cities of Catania, Messina and Palermo; they collect an overall amount of ~2000 cores and a storage capacity of >250 TB.

Figure 1: The sites of COMETA e-Infrastructure.

3 INFINIBAND NET LAYER

Usually, local network connection among the nodes of a Grid site is provided by Ethernet links at 1 Gb/s, since the communication among nodes is not a bottleneck for many applications performing on the Grid. Currently, the majority of the Grid applications are trivially parallel so they can be split in different not communicating chunks. As a result, the bandwidth provided by the GigaEthernet link is sufficient for most types of applications in order to run these jobs efficiently. However, an efficient execution of not trivial parallel jobs requires much faster exchanges of short messages instead of sustained communications. So, a low latency is required. This is the reason why each processor in the Sicilian e-Infrastructure is equipped with two network cards, one for the usual Ethernet connection devoted to Grid services and normal (non-parallel) jobs, and the other one for the InfiniBand connection, reserved to parallel computing only.

The impact of InfiniBand on the programs is noticeable. The latency of communication drops from the Ethernet value of about 50 μs down to 1-2 μs. Optimized parallel programs exploit this feature resulting in a much shorter execution time. The advantage of using the low-latency communication protocol is more sensitive the more nodes are allocated for the computation. For parallel computations involving several tens of cores and more, the adoption of the InfiniBand net layer is highly recommendable. In many cases, the optimization of net communication is a challenging problem, particularly for those parallel codes that have been adapted from different architectures, such as shared memory, to a distributed computing environment. In such cases, the InfiniBand net layer

is very effective too (see also the discussion about MPIGranularity in Section 5).

4 SCHEDULING POLICY

The coexistence of several heterogeneous jobs running on the same infrastructure is a major difference between Grids and dedicated clusters. Particularly, the execution of long lasting and multi-core parallel codes, also called HPC applications, requires an efficient resource assignment. So, the scheduling policy has to implement the best achievable trade-off among the different requirements. Each site has different queues dedicated to short, long and infinite jobs with respective longest durations of 15 minutes, 12 hours and 21 days (lifetime of the longest proxies). The priority is given to the short jobs in order to keep the turnaround time a reasonably small fraction of the execution time. However, this usual scheduling policy is not flexible enough as parallel jobs requiring hundreds of cores (sometimes a significant fraction of all the available resources of a site) would remain scheduled forever. Therefore, the pre-emption and reservations policies have been carefully studied and implemented in order to solve this problem.

Pre-emption puts the incumbent job in a suspended state and assigns the resource to the incoming job. This policy is acceptable only if the latter is short, so it has been implemented for emergency jobs (related to volcanic surveillance and other civil protection applications).

Common parallel jobs reserve the needed resources as soon as they become available. Provided that the amount of short jobs is higher compared to the others, their turnover is fast enough to keep this policy pretty valid. Figure 2 summarizes the scheduling policies. Scheduling is largely dependent on the amount of jobs on each queue and the adopted policy must be updated when significant changes occur in the job distribution.

This issue also witnesses the importance of the overall infrastructure monitoring that has to collect periodic statistic information about job distribution in the various queues.

Gustav is a CPU accounting tool developed by INFN, COMETA and KISTI institute in South Korea. Gustav collects accounting records from resources and publishes them to a centralised relational database that can be queried through a web interface (gustav.consorzio-cometa.it).
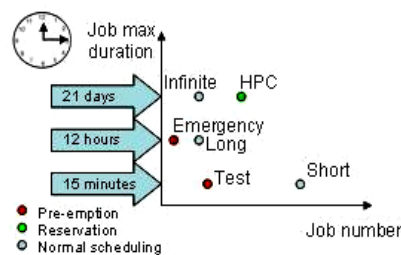

Figure 2: Main queues durations and populations.

5 UPGRADES OF THE GLITE MIDDLEWARE

The gLite middleware supports the Message Passing Interface (MPI) as its only libraries for parallel computing. The complexity of the porting procedure is a major factor against the adoption of the Grid paradigm for HPC. For this reason new wrappers have been developed to support each available combination of compilers, net layers and library versions. The MPIType tag added to the gLite Job Description Language (JDL) allows users to simply select the proper value to run their parallel code in the proper HW/SW environment.

MPIGranularity tag reserves the desired number of cores on the same physical processor for the incoming job. Thus. the communication over the net, a usual bottleneck in such situations, is reduced. Figure 3 shows how the impact of this technique on the execution time is higher compared to that of the InfiniBand net layer.
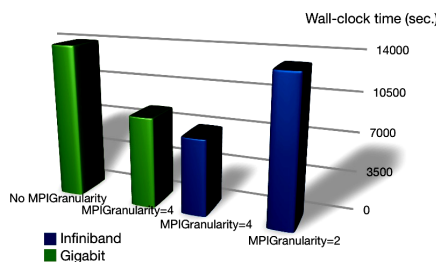
Figure 3: The impact of MPIGranularity and InfiniBand on the wall-clock execution time of a parallel application.

Although gLite has recently extended the MPI support including the so-called "mpi-start" approach, based on a set of scripts able to identify the library requested for each application, the advantages of the developed solution do not allow to move to the official support and some of the ideas were submitted to the MPI Working Group of EGEE

who is evaluating their integration in a future release of the gLite middleware.

6 THE COMETA LICENSE SERVER AND ROBOT CERTIFICATES

The delivery of license files to authorized users in a distributed environment is a non-trivial task. For security reasons, the license server is a single machine for the whole infrastructure, only locally redundant for the sake of service continuity. Thus, the license file has to be delivered to the remote infrastructure sites travelling on a public network. Moreover, solely entitled users must be granted.

The license system developed by COMETA has been designed to be used with the FLEXlm (www.globes.com/support/flexlm_index.htm) free license server that is the most used license management system for commercial software. The "floating" license mechanism has been adapted to the Grid distributed environment, evolving into GridLM. The secured communication channel created from the remote Computing Element (CE) to the license server allows to first identify the user by his/her certificate, check the number of granted licenses and then issue the license file.

Figure 4 illustrates the schema of the license server mechanism available on the COMETA e-Infrastructure.

The team of Grid developers is currently extending the license server mechanism to cope with "robot" certificates as well. They are a novel feature that allows running a program without any personal certificate. The user can store this special certificate on a smart card and insert it into a personal computer running the user interface (usually a web portal). This is enough to authenticate and launch a program on the Grid. This solution fits the requests from large user communities that want to simplify the approach to Grid computing. The accounting system is another key element for a sustainable use of the Grid infrastructures especially when they are open to commercial exploitation. gLite offers various tools for resource metering. PI2S2 project developers extended the measuring to the disk storage with SAGE (https://forge.eu-eela.eu/forum/forum.php?forum_id=31).

This tool computes the amount of disk energy used, i.e. the amount of space integrated over the time of occupation.

Figure 4: The Schema of the COMETA License Server Mechanism.

After a first version, tailored for DPM (www.gridpp.ac.uk/wiki/Disk_Pool_Manager) based devices, a second version was developed that is adaptable to a generic disk storage device. SAGE implements a good trade-off between precision and computing demand of the metering tool. A dedicated web portal called Sight-on-SAGE has been developed as well to allow an easy access to the accounting data both for users and VO managers.

7 SECURE STORAGE

One of the main benefits of Grid infrastructures is the possibility to use distributed storage space. A community might like to use Storage Elements (SE) owned by an external organization to delegate the management of these machines and to avoid to buy specialized hardware and to hire specialized personnel.

In this way the community could rent the storage space as needed and minimize both human and hardware costs.

In the case of confidential data, this scenario is not feasible. Indeed, the community should satisfy strong privacy requirements, for example when it has to manage medical or financial data. A mechanism to prevent the administrator of the machine accessing the data is required to store the confidential data in a storage element managed by an external organization.

The gLite middleware provides the same security infrastructure for all its services (using X.509 infrastructure to authenticate the users and the VOMS attributes to authorize the users).

However, data are stored in a clear format. The storage element administrator can in principle access them bypassing the Grid security infrastructure. This is known as the insider abuse problem.

The Secure Storage Service developed for the gLite middleware provides users with a set of tools to store in a secure way and in an encrypted format confidential data on SEs solving the insider abuse problem. Data stored trough the tools provided are accessible and readable only by authorized users.

The Secure Storage Service (www.ias07.org/) has been designed to be integrated in the gLite middleware; it is made up by the following components:

- Command Line Applications: commands integrated in the gLite User Interface to encrypt & upload and decrypt & download files on the Storage Elements;
- Application Program Interface (API): allowing the developer to write programs able to manage confidential data using the Secure Storage Service;
- Keystore: a new Grid element used to store and retrieve users' keys in a secure way;
- Secure Storage Framework: a service component, internally used by the other components; it provides encryption & decryption functions and other utility functions; it takes care of interaction with the gLite Data Management System.

8 WATCHDOG, VISUAL GRID AND RECURSIVE CATALOGUE INTERACTION

Many complex jobs often need long execution times. Such a long period increases the probability to have errors due to either infrastructure or program faults. Even a brief network interruption may cause the job to fail. On the other hand, if the job is performing an erroneous computation due to wrong input data or application bugs (many applications are undergoing a continuous development and many executions are needed to produce bug-free codes) the user can verify the output only after the execution. This leads to a consequent waste of time and resources, even though many problems could have been revealed a few moments after they have occurred during the run. Due to the intrinsic complexity of parallel jobs and their long durations, they require constant job monitoring and checkpoint features.

Although current gLite middleware offers the perusal job technique for job inspection, COMETA e-Infrastructure offers two more sophisticated tools for job monitoring: Watchdog and VisualGrid.

As the evolution of the job is detectable by inspecting the files produced by the job in the working directory, it is useful to have them copied on the Grid file catalogue at regular intervals, where the user can access them. The Watchdog utility performs such full or incremental backup by a customizable script also allowing the storage on the AMGA (amga.web.cern.ch/amga) metadata catalogue.

The VisualGrid tool (Andronico, 2010) allows the encoding of images produced by the job into a video that is immediately streamed to a public IP address. The result is a powerful tool for demonstrations but also for a direct, visual, control of the job work flow.

The tool has been tested with the FLUENT (www.fluent.com/) commercial application and its simulation of the Marmore falls in Italy (see Figure 5).
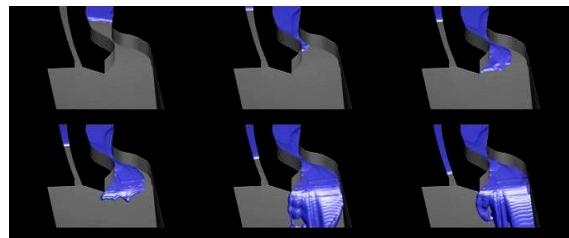

Figure 5: Some frames of the sequence produced by FLUENT and streamed by the VisualGrid tool.

Other tools developed by COMETA help users in their interaction with storage resources. Sometimes, MPI jobs write many files organised in (sub-) directories. The current gLite middleware forces the user to download them one-by-one unless he/she produces an automatic script. Similarly, there is no command for bulk upload or to delete all the files in a directory of the data catalogue. As these operations become more frequent with MPI jobs and the users may not want to write their own scripts for such standard operations, some new tools have been developed for recursive interaction with the Grid data catalogue. Their explanation is available on the web (https://grid.ct.infn.it/twiki/bin/view/PI2S2/WikiConsorzioCometa). The three basic commands are listed in the following: `lcg-rec-cr:` uploads all the files of a local directory on a Storage Element (SE) and registers the entry on the LFC catalogue, copying the structure of the sub-directories and its content as well; `lcg-rec-cp:` performs the recursive download of an entire directory together with its sub-directories; and `lcg-rec-del:`

removes all the entries of a directory from the catalogue and the corresponding files form the SE.

9 PORTING PROCEDURES AND EXECUTION PERFORMANCES

The porting procedure consists in bringing a program already running on a common dedicated platform to run in a Grid environment. The procedure is usually straightforward for simple jobs but it may become rather complicated for complex parallel jobs. This stresses the importance of the computational schema, i.e., the strategy followed to exploit the computing opportunities offered by the Grid. Nevertheless, as the knowledge about the Grid mechanisms and tools is greatly useful for job optimisation, a Grid expert usually helps the user during the porting process of his/her application.

This strategy was adopted in several cases. For instance, the FLUENT parallel implementation required an ad-hoc computing schema as the package is given as a "close box", with no access to the code. Provided that the COMETA HW was supported, FLUENT has been "embedded" into the infrastructure.
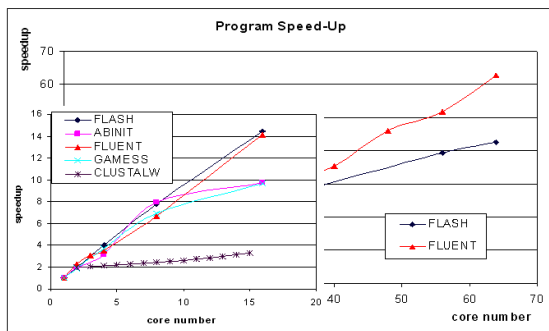
Figure 6: Speed-up of some parallel applications on the COMETA e-Infrastructure.

As shown in Figure 6, the result has been a pretty sensitive speed-up in the execution time comparable with similar dedicated clusters and confirmed by some other parallel applications like FLASH (Orlando, 2007), CLUSTALW (Lombardo, 2010), ABINIT and GAMESS (www.pi2s2.it/applications/), belonging to other scientific domains.

10 CONCLUSIONS

The Sicilian e-Infrastructure has been fully adapted to real business applications. Future developments concern the development of further tools supporting the exploitation of the infrastructure for business purposes. The targeted audiences range from small/medium enterprises to large public bodies such as health care institutions. Interested readers can contact the first author at marcello.iacono@ct.infn.it to get detailed information about the COMETA activities.

ACKNOWLEDGEMENTS

REFERENCES

Andronico, G., Barbera, R., Iacono Manno, M., La Rocca, G., 2010, VisualGrid: On-Line Video Streaming for Application Control and Demonstration, *Proceedings of the Final Workshop of Grid Projects "PON Ricerca 2000-2006 – Avviso 1575"*, ISBN 978-88-95892-00-3, pp. 128-131, published in Catania by the COMETA Consortium

Orlando, S., Peres, G., Reale, F., Bocchino, F., Sacco, G.G., Miceli, M., Bonito, R., Pagano, P., Argiroffi, C., Yelenina, T., 2007, High Performance Computing on the COMETA Grid Infrastructure, *Proceedings of the Grid Open Days at the University of Palermo*, ISBN 978-88-95892-00-9, pp. 181-188, published in Catania by the COMETA Consortium

Lombardo, A., Lanzalone, G., Muoio, A., Barbera, R., Iacono Manno, M., 2007, Performance of ViralPack applications on PI2S2 Grid infrastructure, *Proceedings of the Final Workshop of Grid Projects "PON Ricerca 2000-2006 – Avviso 1575"*, ISBN 978-88-95892-00-3, pp. 446-451, published in Catania by the COMETA Consortium

POSTER

SOA BASED E-BUSINESS SYSTEMS DESIGN

Edzus Zeiris and Maris Ziema

Faculty of Computer Science and Information Technology, Riga Technical University, Meza 1/3, LV-1048, Riga, Latvia
{edzus.zeiris, maris.ziema}@zzdats.lv

Abstract: The development of e-business promotes the creation of new e-services; consequently, ways of fast and quality designing of e-services are required. When developing e-services in SOA, it is very important to build the architecture of e-business system that makes the e-service compliant with all and any quality criteria (QoS) specified for it, which would expand its usability; furthermore, it is necessary to implement any changes swiftly and in good quality to be able to adjust to the rapidly changing business environment. This means that effective design methods should be used in creating e-business systems and e-services, which would ensure the building of an acceptable e-business system architecture. A drawback of the existing methods is the subjective opinion of the system's architect, and that may not always lead to the best solution. Therefore, it is possible to apply the Quality Attributes Driven Design method for web services that is based on the use of formal optimisation methods. Initially the e-service is described as an algorithm graph, and by segmenting its vertices in all possible ways the web service graphs are obtained. The segmentation of the algorithm graph means that all the possible solutions that can affect the quality of the e-service system architecture are dealt with. Using multi-criteria optimisation, a Pareto optimality set is obtained from all the web service graphs. Web service graphs of the obtained Pareto optimality set can serve as the basis for selecting an acceptable e-business system architecture.

1 INTRODUCTION

The global dynamic advancement of information technologies has reached a level where there is high demand for simple and effective means for receiving, processing, storing and exchanging information. Information technologies have become widely available to the public, and it creates demand for e-services that facilitate the advancement of e-business. There is a particular need for e-services in the public sector (government and local government spheres), where there is a very wide range of customers and the number of services is high. More and more government and local government institutions offer their customers services that can be delivered electronically. Compared to government and local government institutions, the advancement of e-services is faster in the private sector. The reason for the advancement of the e-business environment in the private sector could be cost-effectiveness. In the long run, the use of e-services saves resources, which means profit for private commercial structures. Areas where commercial

e-services are most advanced are internet banking, internet shops, information publishing, communications etc. In the public sector, the implementation of e-services is hindered by the bureaucracy and the existing legislation procedures that require a lot of work and time to make them electronic.

Undoubtedly, the availability, range and variety of e-services are required for the convenience of the customers (service users) and to save resources of the service providers. If the range of e-services becomes wider, the clients will benefit from saving their time and resources; therefore, it is important to expand the range of e-services both in the public and private sectors. Development costs are one of the factors that hinder the implementation of e-services. Another factor, which, in terms of making services electronic, affects more the public sector, is the requirements to change the legislation and procedures. Also, it is essential that simultaneous processing of a number of communication channels has to be considered when making services electronic in the public sector. Since services of the public sector must be available to all customers, it
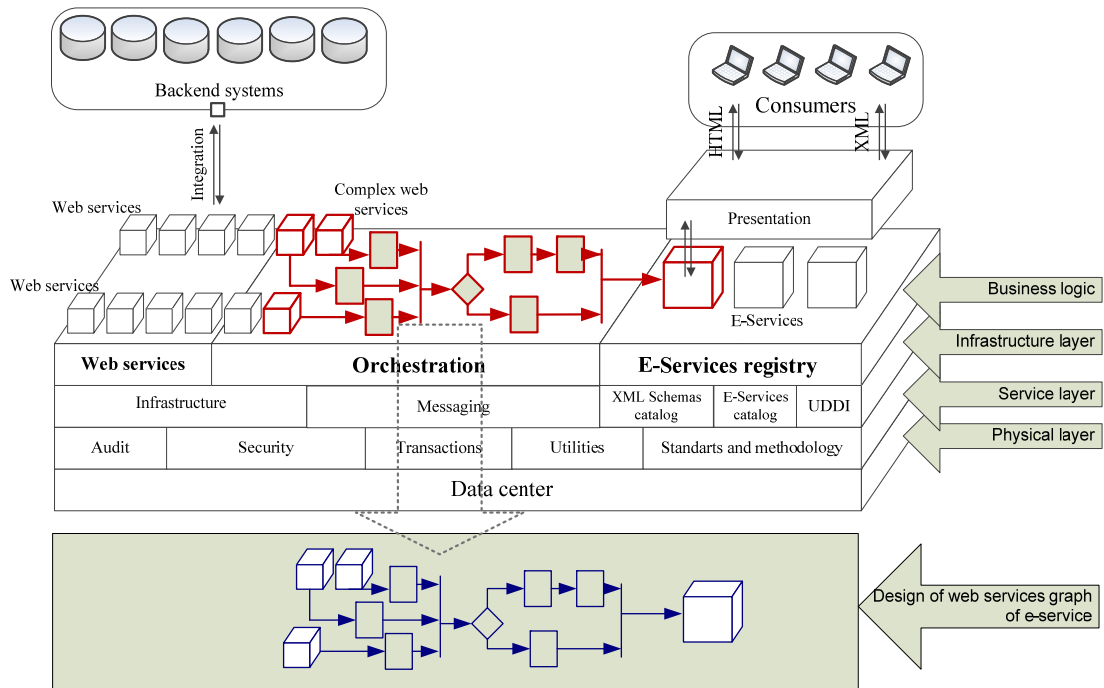
Figure 1: e-Business system architecture

has to be ensured that they can receive services on the Internet, in presence, by phone, mail, e-mail etc.

In the development of e-services, another problem is the rapidly changing business processes that require relevant modifications of IT systems, inter alia e-services. To develop e-service systems in such a dynamic environment, a fast and quality design method that makes it possible to adapt to the business processes and satisfies all the requirements specified for the e-service system is required. It is of great importance to meet the quality of service (QoS) requirements that is the basis for developing a successful e-business.

Currently, the provision of services in the form of web services is advancing at a rapid pace. This is due to the fact that the business-related functionality is offered as ready-made applications (services). Usually a number of web services are combined in e-services. To create an e-service, it is necessary to strictly define service interfaces and their functionality and to organise the performance of the e-service in certain succession. It means that e-business systems are created on the basis of SOA.

Several quality criteria can be specified for the e-business system, and its architecture must comply with them: development costs, maintenance costs, integrity, scalability, security etc. To evaluate the quality of system architecture, ISO 9126 Standard for Software Engineering-Product Quality

Assessment and its quality metrics and guidelines for the use thereof can be used.

Quality criteria are usually mutually controversial; therefore, to be able to comply with all the quality criteria set for the system, it is necessary to design a system architecture that, to a larger or lesser extent, meets all the requirements. In such cases, it is not enough to use functionality-based or object-oriented design approach. It is necessary to apply methods that solve architecture design quality problem.

To ensure the creation of an acceptable e-business system architecture and to be able to adjust e-services to the constantly changing business environment and its requirements, it is possible to employ formal optimisation methods in the designing of the system architecture.

2 E-BUSINESS SYSTEM ARCHITECTURE

On the whole, an e-business architecture, within which e-services are designed, is rather complex and consists of a number of parts. It includes all the components, conditions and mutual links required to design e-services. Fig.1 shows how the systems and system components used in the service are combined into a single e-business system architecture.

Fig. 1. shows how the systems and system components used in the service are combined into unified e–service system architecture. For every data object that is required in the implementation of e–service, an XML Schemas Catalog has to be developed. Data call from the relevant functional system is done by means of the web services. When web service calls are done, also metadata that describe the request are sent. Also any information that is required for the audit trails is sent together with the metadata. Web services can be distributed into two groups: simple and complex. Complex web services in fact are logical combinations of several simple web services that result from the process integration requirements; a combination may comprise simple services of one or several independent functional systems. Complex web services can be executed by using a BPEL processor. A BPEL processor is used as the orchestration (integration) environment for the e–service's web services. Portals, one–stop agency applications etc ensure the delivery of e–services to the users. E–service entry forms, stop points, information on payments and execution results are transferred through HTML or HML pages that can be used in the portal to implement the service using the XSLT transformation. Web service and e–service holders, i.e. institution specialists and system administrators who are responsible for the maintenance and development of web services and e–services, must have a possibility to intercommunicate on various issues connected with the execution and advancement of web services and e–services. Also, asynchronous e–services have to be executed. Messaging systems are designated for this purpose. The messaging system enables working with text messages and work tasks. The application is inter-integrated with the e–services register, from which it receives data on XML schemas, web services and e–services. On the other hand, the messaging application is a client of the orchestrations as messages on the execution of web services and e–services are received from them. Data on all XML schemas, web services and e–services are entered in a single register called E–services Register. The register keeps all the versions of schemas, services and e–services so that this information is accessible to anyone who is engaged in the development and advancement of e–services.

The main problems in the e-business system architecture are related to the e-service's data standardization, audit, the creation of a web service catalog (UDDI), the execution of asynchronous e-services, security and the development of e-service orchestration that is connected with web service design. If e-business environment components that ensure quality e-services are once created, they can be later used repeatedly, whereas web services of particular e-services and their orchestration must be developed for every e-service individually. Furthermore, the development of web services and their orchestration significantly affects the e-service's performance metrics and, consequently, the overall quality of e-service (QoS). To design web services and their orchestration, the Quality Attributes Driven Web Services Design Method, which ensures swift and quality designing, can be used.

3 QUALITY ATRIBUTES DRIVEN WEB SERVICES DESIGN METHOD

The offered e – business system architecture design method consists of the following steps (Fig. 2.):

1. Initially it is necessary to create the e-service algorithm graph G that describes the e-service to be designed;
2. In the e-service algorithm graph, it is necessary to determine the possible restrictions for inter-segmentation of vertexes;
3. By recursively segmenting the e-service algorithm graph, the web service graphs $X = \{G'\}$ are obtained that are then used as the basis for the development of the e-service system architecture; The e-service algorithm graph is assumed as the initial web service graph;
4. The numerical vales of quality metrics of the obtained web service graphs are calculated;
5. From all the web service graphs, the Pareto optimality set $P = \{G'^*\}$ is obtained;
6. Web service graphs of the obtained set P can serve as the basis for selecting an acceptable e-business system architecture. The web service graphs that are contained in the obtained Pareto optimality set can be designed in detail and implemented in the e-business system architecture;
7. In most cases, the Pareto optimality set P contains more than one solution. If the obtained set is sufficiently small, the selection of web service graphs from the set can be left to the system designer. However, if there are many possible solutions, then in order to select a particular web service graph, the criteria can be decreased or combined with another optimisation or design method.
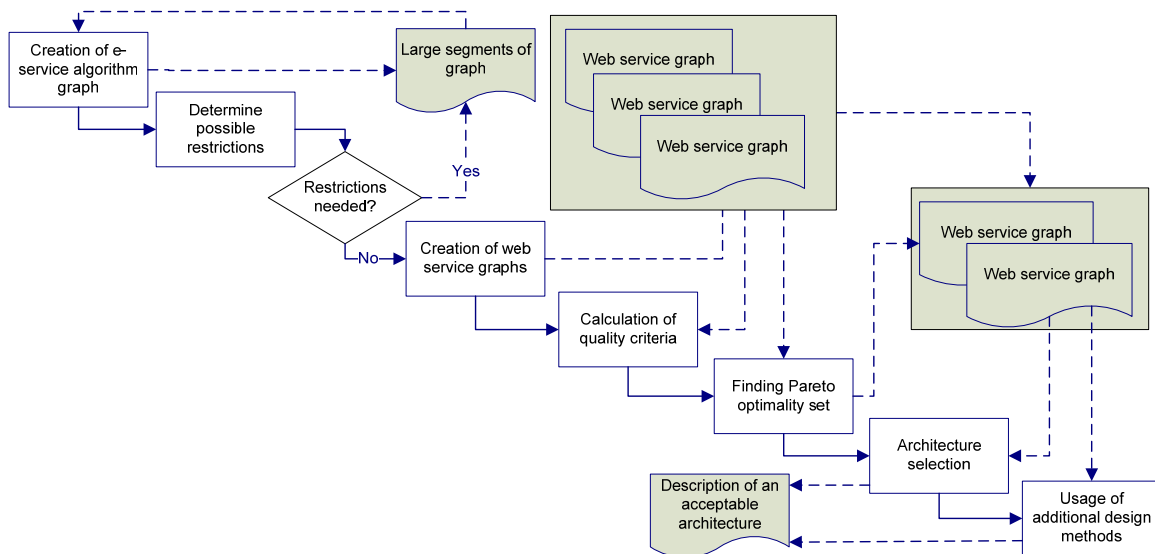
Figure 2: Quality attributes driven web services design method.

3.1 e-Service Algorithm Graph

The e-service algorithm graph is defined as an oriented graph $G = (S, L)$, where $S = \{s_1, s_2, s_3, \cdots s_n\}$ is a final set – graph vertexes, which, to their substance, are the executable operations of the e-service algorithm, and $L \subset S \times S$ are edges. The edge $l = (s_j, s_k)$ in the graph means that in the execution of the e-service algorithm the execution of the operation s_j is followed by the execution of the operation s_j. The edges in the e-service algorithm graph show the information flow (Fig. 3). When creating the e-service algorithm graph that can be used as the basis for developing the e-business system architecture, some of the restrictions related to SOA have to be considered:

1. Every algorithm graph vertex s_i must perform an independent executable activity. It means that the vertex operates in one transaction and is not connected with the algorithms executed in other vertexes. This condition is connected with high cohesion and minimal dependence.

2. Every vertex must contain at least one executable operation $M_{s_i} = \{m_{s_i}^0, m_{s_i}^1, m_{s_i}^2, \cdots m_{s_i}^g\}$. In practice, this is related to the algorithm implementation methods.

3. 3. The operations that repeat during the execution of the algorithm may not be implemented in the graph as various vertexes $\forall s_i, s_j, \ s_i \neq s_j, \ s_i \in S, \ s_j \in S, \ M_{s_i} \cap M_{s_j} = \emptyset$.

   It is required in order to ensure high cohesion and to exclude, from the very beginning, the processing of unnecessary variants.
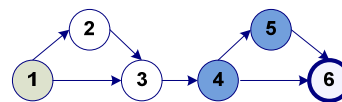
Figure 3: Example of e–service algorithm graph.

3.2 Web Services Graph

The basis for selecting the e-business system architecture is the selection of web services. It means that the e-service algorithm has to be implemented as web services in order to design the e-business system architecture. The algorithm graph can be realised as a web service graph in several ways: from realising every algorithm vertex as an individual web service to realising all the algorithm vertexes as one web service. It means that various e-service algorithm graph segments can be realised as web services, in this way changing the e-service architecture and affecting various e-service execution metrics (Fig. 4.).
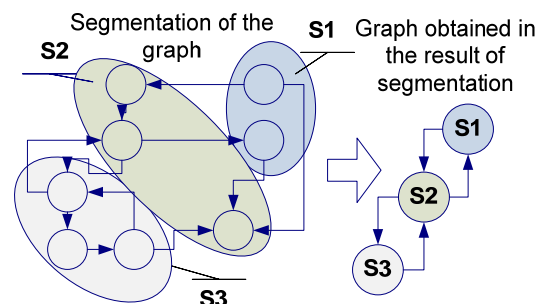

Figure 4: Segmentation of the graph.

To segment the graph and describe it formally, we will deal with the graph G as the depiction Γ in the vertexes set S, attributing to every vertex a sub-set $(s_i \subset S)$, which actually consists of the vertexes that can be reached from the vertex s_i. Γs_i is the edges coming out from the vertex s_i, and $\Gamma^{-1} s_i$ is the edges going into s_i. For such e-service algorithm graphs, there is at least one vertex $s_0 (s_0 \in S)$, which none of the edges $\Gamma^{-1} s_0 = \emptyset$, which we will call the origin of the e-service algorithm, enters into; similarly, there exists at least one vertex $s_r (s_r \in S)$, from which none of the edges $\Gamma s_r = \emptyset$, which we will call the result of the e-service algorithm, is going out. To find all the possible web service graphs from which the one to be implemented in the e-business system architecture is selected, initially the given algorithm graph G is assumed as the first possible web service graph. All other web service graphs are obtained by segmenting the initial web service graph, which concurs with the algorithm graph. Since several segmentations are possible, the set $X = \{G'\}$ that contains all the possible graphs that are recursively derived from the initial web service graph G is defined. Graph transformations are done by merging the vertexes. The combination s' of the vertexes s_i and s_j is as the combination of the ingoing and outgoing edges of both vertexes. $\Gamma s' = \Gamma s_i \cup \Gamma s_j$ and $\Gamma^{-1} s' = \Gamma^{-1} s_i \cup \Gamma^{-1} s_j$. The set of methods in the new created vertex develops as follows: $M_{s'} = M_{s_i} \cup M_{s_j}$.

In practice, when creating the web service algorithm graph, there are a number of restrictions, e.g. it is necessary to integrate the functionality of various systems in one e-service, or it is required to use already existing web services. Therefore, restrictions (vertex feature $I(s_i)$) can be applied to the algorithm graph, which is assumed as the original web service graph. In the given example (Fig. 3.) vertexes 4 and 5 are created in another system, therefore they are marked in another colour. Consequently, $I(s_1) = I(s_2) = I(s_3) = I(s_6) = 0$ and $I(s_4) = I(s_5) = 1$. In this case, the combination of vertexes s' only is possible if $I(s_i) = I(s_j)$. Considering that the algorithm graph that initially is assumed as the web service graph has attributed features, the combination of vertexes s' only is possible if $I(s_i) = I(s_j)$.

Restrictions must be applied to the combining of algorithm graph vertexes and the development of web services also to be able to process large algorithm graphs and find all the possible graph vertexes segments. If there are no restrictions and the algorithm graph is complex, segments have to be extracted in the e-service graph, and then every segment is treated as an individual design task. The division of graphs is related to computation tasks, where it is necessary to divide the graph in two (or more) large parts, minimising the number of edges that cross the split.

Obtaining all the possible algorithm graph segments is an NP–complete problem, where all the possible web service graphs only may be obtained by executing a recursive algorithm.

3.3 Pareto Optimality Set

There are several criteria that have to be taken into account when selecting an acceptable e-business system architecture, and very often they are contradictive. It means that it is necessary to find a web service graph in the set X (which has been formed after segmenting the algorithm graph) that is the best (or at least not worse) according to the specified quality criteria and on the basis of which an acceptable e-business system architecture can be build. It is very essential to deal with all the criteria simultaneously, as it is impossible to determine which of them is more important, and also, they are not mutually comparable. The task of finding the web service graph can be reduced to a multi-criteria optimisation task. Multi-criteria optimisation means finding the Pareto optimality in the sets, which can be formally defined as follows:

$$Q(X) \to \min_{X \in \Omega} \to P \qquad (1)$$

$Q(X) = \{q_1(X), q_2(X), \cdots q_N(X)\}$ are criteria to be minimised; Ω – X definition area. The solution $X_i \in \Omega$ will be called a Pareto optimality if and only if $X_j \in \Omega$ does not exist such that

$$q_i(X_j) \leq q_i(X_i) \qquad (2)$$

for all $i = \{1, 2, \cdots N\}$, where at least one is a strong inequality. In other words, for any value $X_i \in P$ in the set Ω no better than the selected value can be found according to every criterion to be optimised. Depicting the Pareto optimality set graphically by two criteria to be optimised, they are all solutions that are nearest to the origin (Fig. 5.)

To be able to mutually compare and view simultaneously various criteria to be minimised, they must be normalised.

In the selection of the web service graph, the application of the Pareto optimality set is done as follows. An e-service algorithm graph is given. Initially it is assumed as the web service graph G.
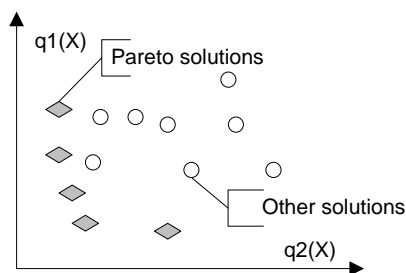
The task is to find the web service graph set

Figure 5: Pareto optimality set.

$P = \{G'^*\}$, which is defined in the area Ω (i.e. all possible algorithm graph segment combinations). The criteria Q, which are the quality criteria specified for the system, are minimised. To determine the system quality criteria, it is possible to use the following methods: scenario-based, simulation-based, mathematical model-based and experience-based.

4 METHOD ANALYSYS AND PRACTICAL USAGE

The application of the Quality Attributes Driven Web Services Design Method has been analysed by using a number of study examples; also, e-business system architecture designs obtained by using this design method and classical object-oriented or functionality-based design methods have been compared. E-services developed for Latvian government and local governments in the portals www.latvija.lv, www.riga.lv and www.epakalpojumi.lv have been used to analyse and approbate the method. The method analysis and the offered e-business system architecture are based on VISS, Latvian Government Information System Integrator, which is an environment for e-services and integration for government and local government services. There are similar e-business and integration systems also in other countries, for example, X-road in Estonia, PSB in Ireland, OIO in Denmark, GovTalk in the UK etc.

Analysing the results obtained from by using the method and comparing them with already implemented e-services, it can be seen that the design method can be used in practice and that it does not have any apparent drawbacks that exclude any essential solutions. The used design methods are different as to the resources that are required for building an acceptable system architecture. It was established that several e-services had been developed iteratively during several years,

improving them constantly to achieve the desired quality, and that several service versions that include improvements in the system architecture had been released. To develop an e-service, a number of experienced system architects had been engaged, which means that resources had been considerably consumed. The offered method achieves the same result by engaging one system architect only and in a comparatively shorter period.

4.1 Design Methods Comparison

To design a complex e-business system, often it is not enough to use one design method, instead, several methods must be combined to meet the required quality of service (QoS). There are several methods that ensure that the required QoS level is achieved, like Attribute Driven Design (ADD), Software Architecture Analysis Method (SAAM), Jan Bosch architecture design method, Architecture Tradeoff Analysis Method (ATAM), Hazard Analysis of Software Architectural Designs (HASARD) etc.

Considering that in the result of using the Quality Attributes Driven Web Services Design Method a Pareto optimality set that contains several results is obtained, then in more complex cases when the system architect is unable to make a selection from the Pareto optimality set, any other design, decision-taking or optimisation method that helps select a solution from the Pareto optimality set can be applied. In this case, the possibility that a local minimum is selected is excluded.

For the system architect to be able to select an acceptable for them design method, a comparison of methods by the following assessment criteria is offered:

1. Multi-Criteria – in selecting the design method, it is important to see to that all the quality criteria specified for the system are dealt with simultaneously;

2. Is the subjective expert position excluded – in many design methods, the selection of solution depends on the system architect, whose subjective opinion influences the system to be designed;

3. Can be estimated in advance – since the development of a system is very time-consuming, it is essential that the quality characteristics can be assessed prior to commencing it;

Table 1: Comparison of architecture design methods.

Method / Comparison Nr.	1.	2.	3.	4.	5.	6.
Functional based	No	No	No	No	Yes	Expert; Methodology
Object-oriented	No	No	No	No	Yes	Expert; Methodology
ADD	Yes	No	Yes	No	Yes	Expert; Methodology
SAAM	Yes	No	Yes	No	Yes	Expert; Methodology
Jan Bosch	Yes	No	Yes	No	Yes	Formulas; Expert; Methodology
ATAM	Yes	Partly, there is more than one expert.	Yes	No	Yes	Experts; Methodology
HASARD	Partly, method is oriented to achieve security requirements.	No	Yes	No	Yes	Expert; Methodology
Quality Attributes Driven Web Services Design Method	Yes	Yes	Yes	Yes	No	Formulas
Combined Quality Attributes Driven Web Services Design Method	No	Partly. To select solution from Pareto compromise set there additional analysis is required.	Yes	Yes	Yes	Formulas; Experts; Methodology

4. Pareto Guaranteed – if the solution is selected from among all the solutions in the Pareto optimality set, it is sure that the selected solution is one of the best possible;

5. Exact Solution – it is important to make sure whether the design method provides one or more solutions;

6. Accuracy of the Solution: expert – the accuracy of the solution is determined by the system architecture's experience; methodology – the accuracy of the solution is determined by complying with the descriptive methodology; formula – the accuracy of the solution depends on the accuracy of the formulae used. As it can be seen from the given comparison (Table 1), there are methods that are multi-criteria and that make it possible to evaluate the quality characteristics of the system architecture prior to development; however, the accuracy of these methods relies only and solely on the expert's experience and the compliance with the methodology. It means that the possibility that a local minimum is selected is not excluded. Another drawback of these methods is the fact that the subjective position of the expert is not excluded, as in most cases only one system architect that takes the decision is engaged in system designing. Considering that successful advancement of computer systems relies upon their quality, any methods that are not multi-

criteria often do not provide the desired result. The application of these methods can cause loss both to the client who ordered the computer system and the developer.

The existing multi-criteria design methods are acceptable in designing large and complex computer systems, which cannot be formally described and for which the mathematic multi-criteria optimisation methods cannot be applied. In such cases the ATAM method, which involves several system design experts, could be very efficient. Also other methods looked at herein could be useful in designing large computer systems if competent experts of the relevant sector and who can take decisions and evaluate the quality characteristics of the computer system to be designed are called in.

To design the e-service system architecture, a design method that provides fast and quality result regardless of the qualification of the involved experts is required. Since the e-service algorithm can be described formally in the form of an algorithm graph, it is possible to use the offered web service design method and the selection of an acceptable e-service system architecture that is based on multi-criteria optimisation. In such and similar cases, it is the multi-criteria optimisation method that should be used to prevent that a local minimum is selected.

Considering that in the result of multi-criteria optimisation a Pareto optimality set that contains several results is obtained, then in more complex cases, when the system architect is unable to make a

selection from the Pareto optimality set, any other design, decision-taking or optimisation method that helps select a solution that is contained in the Pareto optimality set can be applied. In this case, the possibility that a local minimum is selected is excluded.

5 CONCLUSIONS

To meet all the quality requirements set for the e-service, it is important to select an e-business architecture design method that is multi-criteria, as the quality criteria are mutually controversial; otherwise there is a possibility that, while improving one of the quality criteria, another potentially significant condition is not dealt with. If a multi-criteria design method is selected, a compromise among the quality criteria of the e-service is found. A drawback of other design methods is the subjective position of the designer that affects the quality metrics of the designed system and might lead to selecting a local minimum, thus leaving out the best possible solution.

The Quality Attributes Driven Web Services Design Method ensures swift and quality designing of web services, which is the basis for the e-business system architecture and affects the QoS of the e-service. By applying this method, it is possible to optimise the system design costs.

The offered Quality Attributes Driven Web Services Design Method can be used also in other systems that are not complex algorithmically (i.e. algorithm graph can be built) and are based on SOA and web services.

ACKNOWLEDGEMENTS

REFERENCES

Bass L., Clements P., Kazman R. *Software Architecture in Practice 2nd ed.* - Boston: Addison-Wesley , 2003.

Bass L., Klein M., Bachmann F. *Quality Attribute Design Primitives and the Attribute Driven Design Method*// Lecture Notes In Computer Science. Revised Papers from the 4th International Workshop on Software Product-Family Engineering. - London: Springerferlag, 2001. - Nr. 2290 – p.169.-186.

Bosch J. *Design and Use of Software Architectures*. - Harlow: Addison-Wesley , 2000.

Hofmeister C., Nord R., Dilip S. *Applied Software Architecture*. - Massachusetts: Addison-Wesley , 2000.

Kan S. *Metrics and Models in Software Quality Engineering 2nd ed.* - Boston: Addison-Wesley, 2003.

Kazman R., Bass L., Webb M., Abowd G. *SAAM: A Method for Analyzing the Properties of Software Architectures*, International Conference on Software Engineering. Proceedings of the 16th international conference on Software engineering. - Los Alamitos: IEEE Computer Society Press, 1994. –p. 81.-90.

Kazman R., Klein M., Clements P. *ATAM: Method for Architecture Evaluation*. - Pittsburgh: Carnegie Mellon Software Engineering Institute CMU/SEI-2000-TR-004 ESC-TR-2000-004 , 2000.

Losavio F. *Quality Models to Design Software Architecture*, Journal of Object Technology. - 2002. - Nr. Vol. 1, No. 4

Miettinen K. Nonlinear *Multiobjective Optimization*. - Boston: Kluwer Academic Publishers , 1998.

Triantaphyllou E. *Multi-Criteria Decision Making Methods: A Comparative Study*. - Dordrecht: Kluwer academic publishers , 2000.

Zeiris E., Ziema M. *Criterions for Architecture Estimation and Architecture Selection of E – Services System*// Scientific Proceedings of Riga Technical University. Computer Science. Part 5. Volume 27. - Riga: RTU, 2006. – p. 91.-98.

Zeiris E., Ziema M. E-Service *Architecture Selection Based on Multi-criteria Optimization*// Product-Focused Software Process Improvement 8th International Conference, PROFES 2007, Riga, Latvia, July 2007 Proceedings. - Berlin Heidelberg: Springer-Verlag, 2007. - 345.-35

Zeiris E., Ziema M. *Selection of E – Service Systems Architecture*, Scientific Proceedings of Riga Technical University. Computer Science. Part 5. Volume 32. - Riga: RTU, 2007. – p. 99.-107.

Zhu H. Software *Design Methodology From Principles to Architectural Styles*. - Amsterdam: Elsevier Butterworth Heinemann , 2005.

AUTHOR INDEX

AUTHOR INDEX (CONT.)

Proceedings Indexed by:

THOMSON REUTERS
CONFERENCE PROCEEDINGS
CITATIONS INDEX

The Inspec

dblp.uni-trier.de
Computer Science
Bibliography

Ei
(Awaiting confirmation
of indexation)