Spring 4-12-2016

ANALYSING AND EXPLORING DRIFTS IN INNOVATION STREAMS WITHIN OPEN SOURCE (5)

André Alessandro Gomes de Souza
University of Manchester, andre.alessandro@rocketmail.com

"Analysing and Exploring Drifts in Innovation Streams within Open Source"

Submitted by: André Alessandro Gomes de Souza
MA, MSc and Ph. D. Business and Management Student
Supervisors: Dr. Khaleel Malik and Dr. Dimitri Gagliardi
Alliance Manchester Business School
Manchester Institute of Innovation Research – IMP Division

Abstract

This work explores empirically the Apache Hadoop in the context of outbound open innovation (OI) in small and medium-sized enterprises (SMEs) through the lens of innovation streams. The Apache Hadoop is a free and open source (F/OSS) library of codes for distributed computer processing, and it is the industry standard for big data analysis. We are living in the big data age and this research focus on big data analysis digital service platforms. Organisations have radically changed the way they store, manipulate, and create value from information. These data were seen, not very long time ago, as worthless. Businesses are obtaining data from different sources and in diverse formats, and advancing new products and services. Organisations need to explore and exploit niche F/OSS products and services based on outbound OI. Some private sector SMEs are short of tools and require more awareness of the potential benefits of outbound OI for product and service development and the lens of innovation streams offers a multitude of opportunities for analysis. New concepts of value production were brought to light by the notion of OI, including F/OSS. Some private sector businesses lack desorptive capacity, and the proposed conceptual model advances an alternative to the status quo. There is a substantial sum of works on F/OSS, OI and service digital platforms. References to these subjects through the lens of innovation streams in the particular context of the outbound OI in SMEs within the Apache Hadoop appear to be very limited, and there are very few examples of similar studies in this area. Outbound OI is still a major challenge for most firms, some authorities have highlighted the lack of research in the field and expressed the need for complementary studies. Innovation streams are a set of innovations that build upon the current products and services of an organisation, extend that organisation's technical direction, and/or help it diversify into different markets. Outbound OI in F/OSS SMEs' technology spin-offs relates to the innovation streams paradigm in terms of discontinuous innovation. While Michael Tushman and his colleagues have formulated innovation streams in detail, the relation of this framework to the F/OSS outbound OI debate within the Apache Hadoop in SMEs is taken for granted. Many questions regarding this relationship still remain, and this work addresses some of these unanswered issues. This doctoral research endorses the view of an evident limitation in the outbound OI literature, replies to aforementioned calls for more research, and adds to prior analyses by advancing new tools for the comprehension of the role of outbound OI in SMEs. It adds to the emergent body of empirical work on the Apache Hadoop and the current frame of literature on service digital platforms. Its potential findings have implications for both academia and organisations offering big data products and services. Drawing on the qualitative interpretive case study tradition, this research explores theoretical ideas and relates them to the real-world context of Apache Hadoop. This interpretive case study offers suggestions to the following overall research questions: (1) How do innovation streams within the Apache Hadoop evolve from explorative to exploitative and, finally, branch out into new markets? (2) How can we promote and sustain innovation streams within the Apache Hadoop in SMEs, in the context of outbound OI? (3) Can a conceptual model be built? (4) Are these methods adaptable?

1. Introduction

This work empirically explores an open source (FOSS) service digital platform through the analytical lens of innovation streams. It draws, to some extend, on the qualitative interpretative case study tradition, analyses technological circles within the *Apache Hadoop* in small and medium-sized enterprises (SMEs) and extends the innovation streams paradigm to the outbound open innovation (OI) process.

The correlation between economic prosperity and technological shift has being scrutinised extensively and the concept of innovation cycles has been examined across industries from a variety of angles. Due to the remarkable achievements of Michael Tushman and his associates in investigating the drifts in innovation streams in a range of environments, we now have a comprehensive level of understanding in the subject area. These studies have determined and portrayed the diverse phases, as well as concepted patterns within innovation. While Tushman and his fellow academics have detailed innovation streams in many of their works, the relationship of the outbound OI process in FOSS SMEs through the framework of innovation stream is taken for granted. Several problems concerning this affiliation persist and this doctoral research aim attention to some issues ignored so far.

We are living in the big data age and this paper focus on big data analysis digital service platforms. Organisations have radically altered the way they save, manage, and monetise data. Businesses are acquiring data from different sources and in different formats, and developing new products and services. These data were previously cogitated as worthless or too expensive to store (Sammer, 2012). Below in table 1, three definitions of big data by Schneider (2012).

Big Data		
Storing and managing large volumes of data,	Handling diverse data formats	Profiting from these data and new data formats using cutting-edge technology

Table 1: Big Data definitions by Schneider (2012)

Organisations working at big data level have harvested new assets that did not exist at such large scales not long time ago. Businesses have acquired new tools to upgrade their existing services and products or/and to create completely new ones. This work focus on one of the framework suitable to do accomplish such a task.

1.1. Research Questions

- How do innovation streams within the *Apache Hadoop* evolve from explorative to exploitative and, finally, branch out into new markets?
- How can we promote and sustain innovation streams within the *Apache Hadoop* in SMEs, in the context of outbound open innovation (OI)?
- Can a conceptual model be built? Is this models adaptable?

1.2. Main Arguments

This research advances a conceptual model for the development of technology based on F/OSS. Organisations need to explore and exploit niche F/OSS technology products and services based on Outbound OI. Some private sector SMEs are short of tools and require more awareness of the potential benefits of outbound OI for product and service development and the lens of innovation streams offers a multitude of opportunities for analysis.

1.3. Overall Aims

As shown in figure 1, this doctoral proposal discusses elements of technology that draw on three main topics- F/OSS, outbound OI and service digital platforms- through the lens of innovation streams. It focuses on thriving body of literature on OI and service digital platforms, and adds to emergent empirical studies on the *Apache Hadoop*. Based on qualitative interpretive case studies, it suggests a conceptual model for the deeper understanding of how *Apache Hadoop* matures from explorative to exploitative and, later, develop into new products and services. It scrutinises this arrangement and puts forward a conceptual model for academics and practitioners.

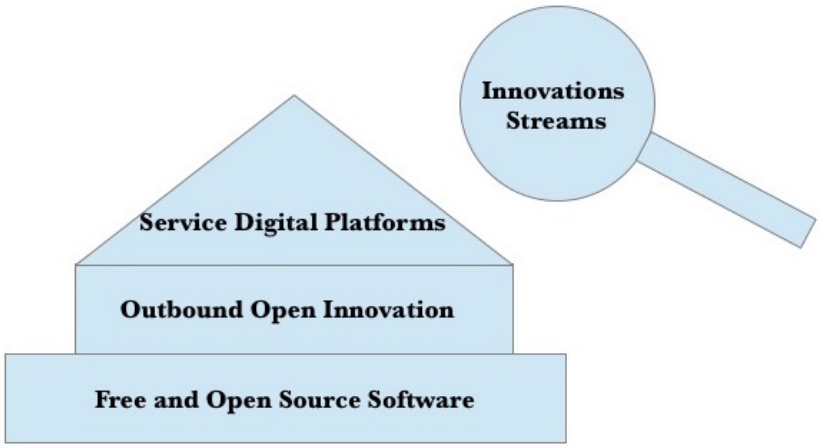


Figure 1: The visualisation of the literature review and the analytical lens

1.4. Specific Objectives

The specific objectives of this thesis are:

- To understand innovation streams within FOSS in the context of outbound OI in depth.
- To identify SMEs pursuing innovation streams within FOSS.
- To investigate the implications of the outbound OI paradigm for sites pursuing innovation streams within FOSS.
- To create a transferable model.

2. Literature Review

2.1 Open Source

F/OSS has radically changed the nature of the creation of value in modern societies. As computer connections increase in speed and reliability, the scale of peer production has gained significance (Benkler, 2006; Baldwin and von Hippel, 2011) and non-professionals and professionals alike have joined forces to produce cultural content such as F/OSS. It is the "quintessential instance" of commons-based peer production (Benkler, 2002) and has turned traditional concepts of software development upside down. Its modular characteristics offer the perfect pillar for the organising of innovation streams. A very encouraging thing is that F/OSS has grown to the point where it has become a major power in today's computing world. F/OSS cannot only seriously challenge the proprietary software industry, but it also represents the forefront of innovation in software development; e.g. Canonical, Oracle and Apache. Joel West and Scott Gallagher (2006) summarised the whole subject in one simple sentence – F/OSS is OI in software. F/OSS exemplifies all the theories discussed in this literature review, and, therefore, it is the object under analysis.

F/OSS represents both a philosophy and a methodology (Stallman, 2002). It gives users freedom and the right to access a library of codes for software development copyrighted under many different open source agreements. It challenges several of the established concepts of software design. Raymond (1999) considered the metaphor of the 'cathedral' versus the 'bazar' model as separating the two very antagonistic means of software development. F/OSS offers a multitude of opportunities to incorporate creative peer networks and gives users access to state-of-the-art technologies. As stated by the Free Software Foundation (FSF), free software is associated with four essential freedoms:

- Freedom to run the software for any purpose
- Access to the source code
- Freedom to make copies and redistribute them
- Freedom to distribute the modified version to others

Recently, the conventional peer-based arrangement highlighted previously has being substituted with:

Sponsored	Is based upon financial injections and/or other kinds of investments from third parties (Capra, 2008).
Industry-led	Is characterised by commercial stakeholders calling the major shots (Hou, 2007; Mens et al., 2008; Merlo et al., 2004; Wermelinger and Yu, 2008).
Industry-involved	Projects are pushed forward by communities but usually have some stakeholders from private or governmental agencies supporting the projects (Capiluppi et al., 2007).

Table 2: New arrangements in F/OSS

This is resulting in a greatly weakened affiliation between communities and organisations, leading to the growth of what Fitzgerald (2006) has formulated as OSS 2.0. OSS 2.0 is described as "the more mainstream and commercially viable form" of F/OSS (Fitzgerald, 2006) or, as Conlon (2011) sums it up, "software designed to automate businesses of a particular type". OSS 2.0 is of major significance for this research.

2.2. Openness

Open Innovation belongs to the extended tradition of studies that shed light on the processes of innovation (Chesbrough and Crowther, 2006). Traditionally, new products were developed within organisational settings and kept safely behind closed doors as a valuable strategic asset. A company's secret was seen as a firm's competitive advantage, and, in order to safeguard it, organisations took advantage of patents and other forms of copyright. In the last decades, it has emerged a common understanding that such rationale is losing its relevance and seen somehow as outdated.

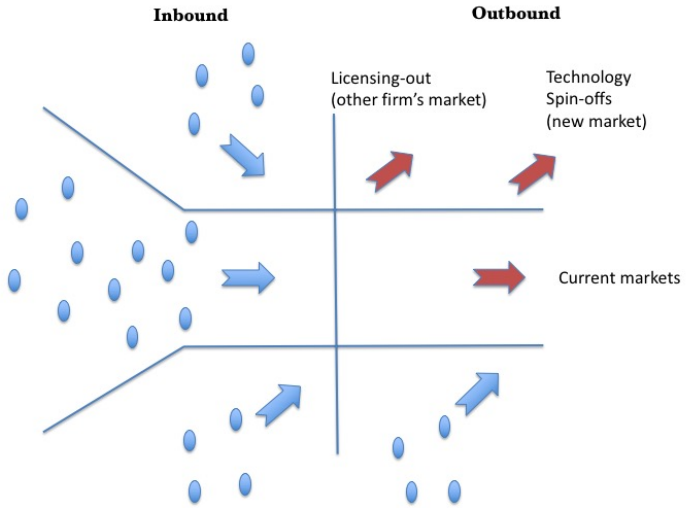


Figure 1: Based upon Chesbrough's funnel (Chesbrough, 2006)

Considering Chesbrough's innovation funnel diagram (2006), as shown above in Figure (fig.) 1, we must make a clear distinction between "inbound" and "outbound" OI. Additionally, table 2 highlights contemporary definition of inbound and outbound OI. A technology can be used in many different ways and it is very unlikely that an organisation can explore/exploit all its countless variations. Businesses should take advantage of secondary markets. Secondary markets widen the means through which cutting-edge technology can be applied and stimulates know-how among market shareholders—essentially, it is a segmentation of the OI process (Chesbrough, 2006). Products and services can enter the market in the outbound OI process in many ways: (1) out-licensing (other firm's markets), (2) Spin-off venture companies (new markets) or (3) the current marketing and sales channels of an organisation itself (Chesbrough, 2011).

Inbound	A leading position is more likely to be achieved by balancing and combining knowledge acquired outside an organisation, with knowledge created internally. This is what some authorities describe as 'open innovation' (Arnand et al., 2002; Lane et al., 2006; Chesbrough and Appleyard, 2007; Boudreau and Lakhani, 2009). Wesley Cohen and Daniel Levinthal (1990) labelled the skills of assimilating internal with external knowledge as the "Absorptive Capacity". It discusses the application of external sources of innovation within an organisation: inward technology transfer or absorptive capacity (Cohen and Levinthal, 1990)
Outbound	It latter considers the usage of peripheral routes to drive development and the commercialisation of an innovation (Chesbrough and Growther, 2006; Lichtenthaler and Ersnt, 2006; Lichtenthaler, 2009; Mortara and Minshall, 2011): outward technology transfer (Lichtenthaler, 2009; van de Vander et al., 2009) or desorptive capacity[1] (Lichtenthaler, 2009; Lichtenthaler and Lichtenthaler, 2009; Ziegler et al., 2013).

Table 3: Inbound and Outbound OI defined

2.3. Service Digital Platforms

Businesses are restructuring themselves due to new specifications for innovative service and product development (Lyytinen and Rose, 2003). These organisations are undergoing operational and cultural changes to adapt their resources and to become more service oriented. These transformations do not happen from one moment to the other; they developed through experiences acquired in the past (Clark, 1985). Innovation is the "recombining or rewrapping" of assets and the more organisations investment in innovation, the more innovations is likely to be commercialised (Arthur, 2009).

When ICTs are combined with other core and peripheral assets in organisations, it allows information to be distributed and reorganized in other sceneries to produce new opportunities for service development and innovation (Lusch and Vargo, 2014). Present attempts to grasp digital infrastructure (Tillson et al., 2010) have highlighted the fruitful features of digital technologies (Henfridsson and Bygstad, 2013), which accelerate service innovation (Yoo et al., 2012).

Service innovation should be considered as developing, shared, vigorous and as knowledge- and information-based, with interaction channels between providers and customers (Miles, 2008). These digital artifacts have been branded as owning an indeterminate rationality (Kallinikos et al., 2013), being intentionally imperfect, uninterruptedly reassembling themselves (Garud and Türtscher, 2008; Zittrain, 2008).

[1] The idea of desorptive capacity was coined to complement the well establish concept of absorptive capacity and characterise the firms' competence to externally exploit knowledge (Lichtenthaler and Lichtenthaler, 2009). Ziegler et al. (2013) have adopted desorptive capacity to describe the firms' ability to externally commercialise their patents.

3. Theoretical Framework

3.1 Innovation Streams

Joseph Schumpeter was one of the first economists to understand the relationship between technical change and economic growth (Dosi, 1982). Carlota Perez (2010) says: "Schumpeter strongly distinguished innovation, seen as the commercial introduction of a new product or a "new combination", from invention, which belongs to the realm of science and technology". It boils down to the simple idea that innovation is a new combination or a new package and according to Tushman et al. (1997) a source of competitive advantage. Innovation is not simply innovation when they vary from one another and are discussed as 'incremental' (Dosi, 1982; Rosenkopf and Nerkar, 2001), 'architectural' (Henderson and Clark, 1990; Baldwin and Clark, 2000) and 'discontinuous' (Dosi, 1982; Tushman and Murmann, 1998; Tushman and Smith, 2002). Below in table 3, the streams of innovations are detailed.

Types of Innovation	What it is	Proposition
Incremental	Incremental innovation is equivalent to normal technological progression, frequently associated to advancement alongside a technological track and expressed by a technological idea (Dosi, 1982)	incremental innovation proposes minimal deviations to the current output, explores the current design, and usually strengthens the supremacy of organisations (Nelson and Winter, 1982; Ettlie et al., 1984; Dewar and Dutton, 1986; Tushman and Anderson, 1986).
Architectural	The concept of architectural innovation is defined as the exploitation of an established product without changing its main components.	Innovations that vary in how the little pieces of a product are coupled, while not changing the underlying parts, are defined as architectural innovations (Henderson and Clark, 1990).
Discontinuous	In addition to exploring and exploiting established technologies sites must attempt to branch out into different markets - to put it simply, businesses must pull strings in opposite directions (Abertnathy and Clark, 1985; Eisenhardt and Tabrizi, 1995; Teece and Pisano, 1994; Tushman and O'Reilly, 1997).	Organisations need to explore new things and apply technologies in new ways in order to remain innovative. Discontinuous innovations rest on a distinctive array of engineering and scientific fundaments and usually advance a brand new market and potential new appliance of a technology (Dess and Beard, 1984; Ettlie et al. 1984; Dewar and Dutton, 1986).

Table 4: The different streams of innovation defined

4. Research Gap

There is a substantial sum of works on open source, OI and service digital platforms. However, references to these subjects through the lens of innovation streams in the particular context of the outbound OI process in SMEs within the *Apache Hadoop* appear to be very limited, and there are very few examples of similar studies in this area.

According to Hu et al. (2015), "outbound open innovation […] remains a challenge for most firms". Some scholars have highlighted the lack of research in the area of outbound OI and expressed the need for complementary studies (Lichtenthaler and Ernst, 2006; Mortara and Minshall, 2011, Ziegler et al., 2013). This doctoral research endorses the view of an evident limitation in the outbound OI literature, replies to aforementioned calls for more research, and adds to prior analyses by advancing new tools for the comprehension of the role of outbound OI in the context of the *Apache Hadoop* in SMEs through the lens of innovation streams. It also adds to the emergent body of empirical work on the *Apache Hadoop*. Therefore, the potential findings have implications for both academia and organisations offering big data products and services.

5. Methodology

5.1. Qualitative Research in the IS Field

IS research deals with technological change and innovation. It discusses technical, managerial and social activities. It positions itself between engineering and social science, and its significance and tenacity are frequently distrusted[2] by both (Avgerou, 2000). IS research offers wide-ranging debates of epistemological paradigms, including positivism and interpretivism (Fitzgerald and Howcroft, 1998; Jones, 2004; Lee, 1991; Mingers, 2001; Probert, 2001; Russo and Stolterman, 2000; Walsham, 1995; Weber, 2004). Qualitative research has frequently been quoted positively by positivists (Yin, 1994) but there is an appealing counterpart of interpretive case study works (Klein and Myers, 1999; Benbasat et al., 1987; Eisenhardt, 1989; Walsham, 1995).

Due to a shift from a technological to a more managerial and organisational agenda (Benbasat et al., 1987; Myers, 1997), the social inquiries associated with IS have come under the spotlight in recent decades (Walsham, 1995). Qualitative research uses qualitative data, such as interviews, documents and participant observation, in order to understand and explain social phenomena (Benbasat et al., 1987; Eisenhardt, 1989).

Bearing in mind the the area of concern, and the analytical nature of this research, the methodology is qualitative (Edmondson and McManus, 2007). Methods within the qualitative tradition present numerous valuable instruments for the study of IS and have been widely applied in the field (Myers, 1997; Orlikowski and Baroudi, 1991; Benbasat et al., 1987; Lee, 1989; Munford et al., 1985; Smith, 1990; Walsham, 2006). According to Walsham (1995), interpretive case studies are of inestimable significance to IS theory and practice, and interviews are the dominant constituent of most interpretative studies.

[2] There is a tension in regards to the essence of IS research (Lee, 2001; Baskerville and Myers, 2002; Avison and Fitzgerald, 2003). Some scholars advocate that the IS field is in disarray as to what the essential concepts of the field are (Orlikowski and Iacono, 2001; Benbasat and Zmud, 2003) and other academics contend that multiplicity is of inestimable value (Walsham, 2012). This multiplicity of theoretical methods has proposed answers from the extremely technical to more philosophical questions (Avgerou, 2000).

5.2. Interpretive Case Study Research

Interpretive research has received increased acceptance in social sciences (Orlikowski and Baroudi, 1991), is deep-rooted in IS research and applied as a tool in distinct topics and inquiries in the field (Klein and Myers, 1999; Benbasat et al., 1987; Eisenhardt, 1989; Walsham, 1995; Markus, 1983; Suchman, 1987; Zuboff, 1988; Boland and Day, 1989; Orlikowski, 1992; Walsham, 1993). It is well-respected in IS research in organisation and more suitable than positivism for research on organisations (Orlikowski and Baroudi, 1991). When used appropriately, cross-case analysis is a reasonable way (Drake et al., 1998) to highlight diverse features of the objects of study (Eisenhardt, 1991). According to Drake et al. (1998), "Multiple case studies allow cross-case analysis and comparison, and the investigation of a particular phenomenon in diverse settings." Broadbent and Weill (1998) and Cavaye and Cragg (1995) are successful examples of interpretive case studies in IS.

6. Field Work

The field work is divided in two part, will last around six months and is planned to take place from September 2016 in Britain and Brazil. At the first stage, interviews will be conducted in Britain with specialists who cover different aspects of and are involved in product and service development around the *Apache Hadoop* framework. It is crucial for the development of the proposed conceptual model that the researcher have input from different professional perspectives-from technical to a more managerial views. The participants are software developers, architecture developer or have a more managerial role in SMEs developing products and services for *Apache Hadoop*.

The second part is in Brazil. In appendix **B** are the organisation that the researcher will contact from March 2016 to negotiate access to the organisations or to be able to interview some key players in those organisations. The researcher has also heard that *LinkedIn* is also a very interesting channel to get in touch with professional working on the *Apache Hadoop*. In March the research will develop a letter in the form of an e-mail in English and Portuguese so he can send around in order to establish contact with organisations and individuals in Britain and Brazil.

The researcher aims to conduct and analyse 40+ interviews across SMEs in Brazil and Britain. This, in turn, will guarantee originality and sufficient primary data to potentially advance an original piece of research.

7. Data Collection and Analysis

The data for analysis will be drawn two ways: (1) secondary data from industry-led case studies and (2) interviews will be conducted with a panel of international *Apache Hadoop* experts who cover a range of skills within the industry. Please see appendix **A** for the details of collaborators and appendix **C** for industry-led case studies web pages. The interviews will be audio recorded, with permission, and transcripts made. The data gathering and analysis will follow strict research ethics as recommended by the University of Manchester. When the audio recording will not be allowed comprehensive notes will be taken instead. A list of same structured interview questions will be designed between March to June and pilot revised with one to three participants.

These questions will serve as a basis for guiding the interviews and are intended to establish open-ended discussions. While the interviews focus relates primarily to the understanding of the innovation circles within the

Apache Hadoop the researcher will encourage the participants to articulate their thoughts on the overall impact and advantages of the the *Apache Hadoop* service digital platforms for day to day business. The researcher is aware of the challengers regarding the design of such questions and he intends to focus on relevant literature on interview design, follow the advice of his supervisors and other more experienced academics.

In order to validate new insights that may arise during/after the data analysis, the researcher plans to re-interview some key participants. This is also a great opportunity to verify some finding and conclusion.

8. Three Paper Path

As discussed with my supervisors this project considers the publication of three academic journal articles instead of a traditional monograph. The ultimate goal is the put three articles on the pipelines of high quality academic journals. The researcher understands how difficult it is to publish in those high caliber journals and will discuss with his supervisors if other opportunities arises as he goes along with his empirical research. It is of extreme relevance that these three articles can contribute to academic knowledge and the researcher will develop what the supervisors have described as the "glue", binding those articles together in the next four months-from March to the first year review in June 2016.

Paper	Target	Availability	Objective
Literature Review (Theoretical Paper)	International Journal of Management Review	No output on outbound OI available in their catalogue.	An article in the pipeline by July 2016
Empirical academic journal article around innovation	Major innovation academic Journal	Empirical Studies on the Apache Hadoop Framework in such journals are very rare or practically non-existent	An article in the pipeline by August 2017
Empirical academic journal article around information systems	Major Information Systems academic Journal	Empirical Studies on the Apache Hadoop Framework in such journals are very rare or practically non-existent	An article in the pipeline by April 2018

Table 5: Three paper (target, availability and objectives)

Before submitting manuscripts to academic journals, the researcher aims to present his empirical findings at major international IS conferences in 2017-2018. He is targeting conferences such as: R&D Management, Association for Information System (AIS) and International Society for Professional Innovation Management (ISPIM).

9. Risk Management

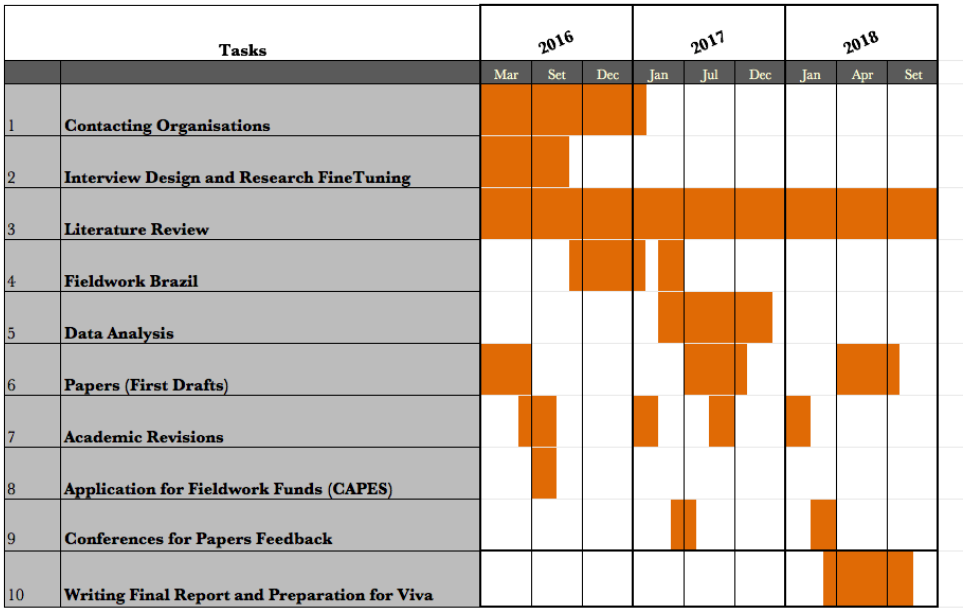
Although the researcher has established a lot of contact with many industry experts it does not necessarily mean that they will have time. It is well known that the agenda of such professionals can change from one moment to the other. A major weakness of this project is that the researcher has not yet established contact with organisations and professional in Brazil. The researcher will scan for big data organisations offering Apache Hadoop products and services in Rio de Janeiro and São Paulo from March 2016, in appendix **B** is a list potential organisations. The researcher is already using *LinkedIn* and have connected with some professional in Brazil. However, he has to contact these organisations thorough a more formal e-mail.

10. Timeline

Doctoral Proposal Timeline

Analysing and Exploring Drifts in Innovation Streams within Open Source

André A. Gomes de Souza
Ph. D. Business and Management Student
Manchester Institute of Innovation Research - IMP Division



Bibliography

Abernathy, W. J., and Clark, K. B. (1985). Innovation: Mapping the Winds of Creative Destruction. *Research Policy*, *14*, 3–22.

Arnand, V., Glick, W. and Manz, C. C. (2002). Thriving on the knowledge of outsiders: Tapping organizational social capital. *Academy of Management Executive*, *16*, 87–101.

Arthur, B. W. (2009). *The nature of technology: What it is and how it evolves.* New York: Free Press.

Avgerou, C. (2000). Information Systems: What sort of science is it? *Omega*, *28*(5), 567–579.

Baldwin, C. Y., and Clark, K. B. (2000). *Design Rule: The Power of Modularity*. Cambridge, MA: MIT Press.

Baldwin, C. and von Hippel, E. A. (2011). Modeling a paradigm shift: From producer innovation to user and open collaboration innovation. *Organization Science*, *22*(6), 1399–1417.

Benbasat, I., Goldstein, D. K. and Mead, M. (1987). The case research in studies of information systems. *MIS Quarterly*, 369–386.

Benkler, Y. (2002). Coase's Penguin, or, Linux and "The nature of the firm." *The Yale Law Journal*, *112*(3), 336–446.

Benkler, Y. (2006). *The Wealth of Networks*. New Haven and London: Yale University Press.

Boland, R. J. and Day, W. F. (1989). The experience of system design: A hermeneutic of organizational action. *Scandinavian Journal of Management*, *5*(2), 87–104.

Boudreau, K. J., and Lakhani, K. R. (2009). How to manage outside innovation. *MIT Sloan Management Review*, *50*.

Cavaye, A. L. M., and Cragg, P. B. (1995). Factors contributing to the success of customer-oriented interorganizational systems. *Journal of Strategic Information Systems*, *4*, 13–30.

Chesbrough, H. (2006). Open innovation: Research a new paradigm for the understanding industrial innovation. In H. Chesbrough, W. Vanhaverbeke, and J. West (Ed.), *Open innovation: Research a new paradigm.* Oxford, United Kingdom: Oxford University Press.

Chesbrough, H. (2011). *Open Services Innovation: Rethink your Business to Grow and Compete in new Area*. San Francisco, California: Jossey-Bass.

Chesbrough, H. and Appleyward, M. (2007). Open Innovation and Strategy. *California Management Review*, *50*, 57–76.

Chesbrough, H., and Crowther, A. K. (2006). Beyond high tech: early adopters of open innovation in other industries. *R&D Management*, *36*(3), 229–236.

Clark, K. B. (1985). The interaction of design hierarchies and market concept in technological evolution. *Research Policy*, *14*, 235–251.

Cohen, W., and Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, *35*, 128–152.

Dess, G., and Beard, D. (1984). Dimensions of organizational task environment. *Administrative Science Quarterly*, *32*, 1422–1433.

Dewar, R. D., and Dutton, J. E. (1986). The Adoption of Radical and Incremental Innovations: An Empirical Analysis. *Management Science*, *32*, 1422–1433.

Dosi, G. (1982). Technological paradigms and technological trajectories. *Research Policy*, *11*(3), 147–162.

Drake, P., Shanks, G. and Broadbent, M. (1998). Successfully completing case study research: Combining rigour, relevance and pragmatism. *Information System Journal*, *8*, 273–289.

Edmondson, A. C., and McManus, S. E. (2007). Methodological fit in management field research. *Academy of Management Review*, *32*, 1155–1179.

Eisenhardt, K. M. (1989). Building theories from case study research. *The Academy of Management Review*, *14*(4), 532–550.

Eisenhardt, K. M. (1991). Better stories and better construct: The case for rigor and comparative logic. *Academy of Management Review*, *16*, 620–627.

Eisenhardt, K., and Tabrizi, B. N. (1995). Accelerating Adaptive Processes: Product Innovation in the Global Computer Industry. *Administrative Science Quarterly*, *40*(1).

Ettlie, J. E., Bridges, W. P. and O'Keefe, R. D. (1984). Organization Strategy and Structural Differences for Radical Versus Incremental Innovation. *Management Science*, *30*, 682–695.

Faulkner, P., and Runde, J. (2009). On the identity of technological objects and users innovation in functions. *Academy of Management Review*, *34*(3), 442–462.

Fitzgerald, B. and Howcroft, D. (1998). Competing Dichotomies in IS Research and Possible Strategies for Resolution. In *19th International Conference on Information Systems ICIS 1998* (pp. 155–164). Helsinki, Finland.

Garud, R., Jain, S. and Tuetscher, P. (2008). Incomplete by design and designing for incompleteness. *Organization Studies*, *29*(3), 351–371.

Helfat, C. E., Finkelstein, S., Mitchell, W., Pereraf, M. A., Singh, H., Teece, D. J. and Winter, S. G. (2007). *Dynamic Capabilities: Understanding strategic change in organizations.* Oxford: Blackwell.

Henderson, R., and Clark, K. (1990). Architectural innovation: the reconfiguration of existing product technologies and the failure of existing firms. *Administrative Science Quarterly*, *35*, 9–30.

Henfridsson, O., and Bygstad, B. (2013). The generative mechanisms of digital infrastructure evolution. *MIS Quarterly*, *37*(3), 907–931.

Hu, Y., McNamara, P. and Mcloughlin, D. (2015). Outbound open innovation in bio-pharmaceutical out-licensing. *Technovation*, *35*, 46–58.

Jones, M. R. (2004). Debatable Advice and Inconsistent Evidence: Methodology in Information Systems Research. In J. I. Kaplan, B., Truex III, D., D. Wastell, A. T. Wood-Harper and DeGross (Ed.), *Information Systems Research: Relevant Theory and Informed Practice*. Boston, Massachusetts: Kluwer Academic Publishers.

Kallinikos, J., Aaltonen, A. and Morton, A. (2013). The ambivalent ontology of digital artifacts. *MIS Quarterly*, *37*(2), 357–370.

Klein, H. K., and Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly*, *23*(1), 67–94.

Lane, P., Koka, B. R. and Pathak, S. (2006). The Reification of Absorptive Capacity: A critical Review and Rejuvenation of the Construct. *The Academy of Management Review*, *31*, 833–863.

Lee, A. (1991). Integrating positivist and interpretivist approaches to organizational research. *Organization Science*, *2*, 342–365.

Lee, A. S. (1989). A Scientific methodology for MIS case studies. *MIS Quarterly*, *13*, 33–50.

Lichtenthaler, U., and Ernst, H. (2006). Attitudes to externally organising knowledge management tasks: A review, reconsideration and extension of the NIH syndrome. *R&D Management*, *36*(4), 367–386.

Lichtenthaler, U. (2009). Outbound open innovation and its effects on firms' performance: Examining environmental influences. *R&D Management*, *39*(4), 317–330.

Lichtenthaler, U. and Lichtenthaler, E. (2009). A Capability-Based Framework for Open Innovation: Complementing Absorptive Capacity. *Journal of Management Studies*, *46*(8), 1315–1338.

Lusch, R. F., and Vargo, S. L. (2014). *Service-Dominant Logic: Premises, Perspectives, Possibilities*. New York: Cambridge University Press.

Lyytinen, K., and Rose, G. M. (2003). The Disruptive Nature of Information Technology Innovations: The Case of Internet Computing in Systems Development Organizations. *MIS Quarterly*, *27*(4), 557–596.

Markus, M. L. (1983). Power, politics, and MIS implementation. *Communication of the ACM*, *26*, 430–444.

Miles, I. (2008). Patterns of innovation in service industries. *IBM Systems Journal*, *47*(1), 115–128.

Mingers, J. (2001). Combining IS research methods: towards a pluralist methodology. *Information Systems Research*, *12*(3), 240–259.

Mortara, L., and Minshall, T. (2011). How do large multinational companies implement open innovation? *Technovation*, *31*, 586–597.

Mumford, E., Hirschheim, R.A., Fitzgerald, G. and Wood-Harper, A.T. (1985). *Research Methods in Information Systems.* Amsterdam: North-Holland.

Myers, M. D. (1997). Qualitative Research in Information System. *MIS Quarterly*, *21*(2), 241–242.

Neely, A. (2008). Exploring the financial consequences of the servitisation of manufacturing. *Operations Management Research*, *1*(2).

Nelson, R. R., and Winter, S. G. (1982). *An Evolution Theory of Economic Change.* Cambridge, MA: Belknap Press.

Orlikowski, W. J., and Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, *2*, 1–28.

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization Science*, *3*(3), 398–427.

Perez, C. (2010). Technological Revolution and Techno-economic Paradigm. *Cambridge Journal of Economics*, *34*(1), 185–202.

Probert, S. K. (2001). Contemporary Epistemology and IS Methodology: An Interpretive Framework. In *7th Americas Conference on Information Systems AMCIS 2001*. Boston, MA.

Raymond, E. S. (1999). *The cathedral and the bazaar: Musing on Linux and open source by an accidental revolutionary.* Cambridge: O'Reilly.

Rosenkopf, L., and Nerkar, A. (1999). On the complexity of technological evolution: Exploring coevolution within and across hierarchical levels in optical disk technology. In J. Baum and W. McKelvey (Ed.), *Variations in Organization Science: In Honor of D. T. Campbell.* Thousand Oaks, CA: Sage Publications.

Russo, N. L., and Stoltermann, E. (2000). Exploring the assumptions underlying information systems methodologies. Their impact on past, present and future ISM research. *Information Technology and People*, *13*(4), 313–327.

Sammer, E. (2012). *Hadoop operations*. Cambridge: O'Reilly Media.

Schneider, R. D. (2012). *Hadoop for Dummies*. Mississauga, ON: John Wiley & Sons, Inc.

Smith, J. K. (1990). Alternative research paradigms and the problem of criteria. In E. G. Guba (Ed.), *The Paradigm Dialog*. Newbury Park, CA: Sage.

Suchman, L. (1987). *Plans and Situated Actions: The Problem of Human Machine Communication*. New York: Cambridge University Press.

Tillson, D., Lyytinen, K. and Sørensen, C. (2010). Digital infrastructures: The missing IS research agenda. *Information Systems Research*, *20*(4), 748–759.

Tushman, M. and Murmann, J. (1998). Dominant designs, technologies cycles and organizational outcomes. *Research in Organizational Behavior*, *20*, 231–266.

Tushman, M. and Smith, W. (2005). Managing strategic contradictions: A top management model for managing innovation streams. *Organization Science*, *16*, 522–536.

Tushman, M. L., and O'Reilly, C. (1997). Winning through innovation. *Strategy and Leadership*, *25*(4), 14–19.

Tushman, M., and Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative Science Quarterly*, *31*, 439–465.

Tushman, M., Anderson, P. and O'Reilly, C. (1997). Technology cycles, innovations streams, and ambidextrous organizations: Organization renewal through innovation streams and strategic change. In M. Tushman, Anderson, P. and O'Reilly (Ed.), *Managing Strategic Innovation and Change*. New York: Oxford University Press.

Vandermerwe, S., and Rada, J. (1988). Servitisation of business: Adding value by adding services. *European Journal of Management*, *6*(4), 314–324.

Walsham, G. (1993). *Interpreting Information Systems in Organisations*. Chichester: Wiley.

Walsham, G. (1995). Interpretive Case Studies in IS Research: Nature and Method. *European Journal of Information Systems*, *4*, 74–81.

Walsham, G. (2006). Doing interpretative research. *European Journal of Information Research*, *15*, 320–330.

Weber, R. (2004). The Rhetoric of Positivism Versus Interpretivism. *MIS Quarterly*, *28*(1), iii–xii.

Weill, P., and Broadbent, M. (1998). *How Market Leaders Capitalize on Information Technology.* Cambridge, MA: Harvard Business School Press.

West, J., and Gallagher, S. (2006). Challenges of Open Innovation: The Paradox of Firms Investments in Open –Source Software. *R&D Management,* 36(3), 319-331.

Yin, R. (1994). *Case study research: Design and methods*. Thousand Oaks, CA: Sage Publishing.

Yoo, Y., Boland, R.J., Lyytinen, K. and Majchrzak, A. (2012). Organizing for innovation in the digitized world. *Organization Science*, *23*(5), 1398–1408.

Ziegler, N., Ruether, F., Bader, M. A. and Gassman, O. (2013). Creating value through external intellectual property commercialization: a desorptive capacity view. *Journal of Technology Transfer*, 38, 930–949.

Zittrain, J. (2008). *The Future of the Internet - And How to stop it.* New Haven, CT: Yale University Press.

Zuboff, S. (1988). *In The Age of Smart Machine.* New York: Basic Books.

Appendix

A.

Name	Function	Organisation	Country	Confirmed/Method	Level of Experience
Vij Jadhav	Developer (architecture, data extraction, tools and database)	Capgemini	India (Mumbai)	Yes / Skype	High
Ferdi Güran	Consultant	Nextevolution	Germany (Hamburg)	No / Skype	Medium
Mateusz Parzonka	Associated IT Consultant	MSG System	Germany (Frankfurt)	Yes / Skype	Beginner
Günther Schnack	Field Sales Manager DACH	DataStax	UK (Middlesex)	No / in person	High
Thomas Gregg	Enterprise Sales Manager	DataStax	Germany (Frankfurt)	No / In person	High
Hakan Lofcali	Software Developer	Etecture	Germany (Frankfurt)	No / In person	Medium
Daniel Cohen	Solution Engineer (water-walker)	DataStax	UK (Middlesex)	Yes / in person	Very High
Christopher Reeddijk & Gary Steward	Advisory IT Specialist	ING	Netherlands (Amsterdam)	No / Skype	High
Patrick Callaghan	Solutions Architect/SWAT (water-walker)	DataStax	UK (Middlesex)	Yes / in person	Very High
Peter Evison	Business Development Manager	Cake Solutions	UK (Manchester)	No / First Contact	Not Sure
Arthur von Scala	Trading System Developer	Credit Suisse	Switzerland (Zurich)	YES / in Person	High

B.

Organisation	Contact Person	Position	City	Service
Semantix	Leonardo Dias	Chef Data Officer	São Paulo	Hadoop
Fio Cruz	Valdir Ermida	Works for Fio Cruz Ph. D. Candidate at AMBS	Rio De Janeiro	Cloudera
Mi Montreal Informatica	André Ribeiro	Customer of the organisation (Detran – Rio)	Rio de Janeiro	Not sure
ICX Soluções	Marcos Colnaghi	Infrastructure Pre-Sales	São Paulo	Not sure
EmergiNet	Edgar Nishiyama	CTO/Data Architect/Researcher	São Paulo	Hadoop
Big Data BRasil	Prof. Eduardo Hruschka	Chief Data Scientist	São Paulo	Not sure

C.

Organisation	**Web-address**
Cloudera	http://www.cloudera.com/customers.html
DataStax	http://www.datastax.com/resources/casestudies
Hortonworks	http://hortonworks.com/industry/
MapR Solutions	https://www.mapr.com/resources/white-papers#.Customer
Pivotal	http://pivotal.io/resources/1/case-studies
Teradata	http://www.teradata.co.uk/Resources/Case-studies/?LangType=2057&LangSelect=true