ANALYSING AND EXPLORING DRIFTS IN INNOVATION STREAMS WITHIN OPEN- SOURCE

A Thesis Submitted to the University of Manchester for the Degree of

Doctor of Philosophy (PhD) in the Faculty of Humanities

2020

ANDRÉ A. GOMES DE SOUZA

Manchester Institute of Innovation Research

Alliance Manchester Business School

TABLE OF CONTENTS

This thesis contains 84,786 words including title page, tables and footnotes but excluding references and appendices.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACID	Atomicity, Consistency, Isolation and Durability
ANATEL	Agência National de Telecomunicação
API	Application Programming Interface
AT&T	American Telephone and Telegraph
ARPANET	Advanced Research Projects Agency Networks
A2A	Actor-to-Actor
BCBS	Basel Committee on Banking Supervision
BMF	Business Model Framework
BRP	Business Process Re-engineering
BSD	Berkley Software Distribution
BVRJ	Rio de Janiero Stock Exchange
CGI	Common Gateway Interface
CEO	Chief Executive OIfficer
CTO	Chief Technology Officer
CQL	Cassandra Query Language
F/OSS	Free and Open-Source Software
FSF	Free Software Foundations
GB	Gigabyte
GDPR	General Data Protection Regulation
GE	General Electric
GNU	GNU’s not Unix
GPD	Gross Domestic Product
GPL	General Public License
GSM	Global System for Mobile Communication
HDFS	Hadoop Distributed File System
HTTP	Hypertext Transfer Protocol
IBM	International Business Machines
IBGE	Intituto Brasileiro of Geografia e Estatística
ICT	Information and Communication Technology
IDC	International Data Corporation
IMF	International Monetary Fund
IP	Intellectual Property

IPTV	Internet Protocol Television
IS	Information Systems
ISIC	International Standard of Industrial Classification
IT	Information Technology
JIT	Just-in-Time
G-D	Goods-Dominant
LGPL	Lesser General Public License
LSI	Local Systems of Innovation
MIT	Massachusetts Intitute of Technology
MPL	Mozilla Public License
MVP	Most Valuable Player
M2M	Machine-to-Machine
NIH	Not Invented Here
NSI	National Systems of Innovation
NSO	Not Sold Here
OECD	Organization for Economic Cooperation and Development
OSI	Open Source Initiative
OSS	Open Source Software
PC	Personal Computer
PERL	Practical Extraction and Report Language
P&G	Procter & Gamble
R&D	Research and Development
RSI	Regional Systems of Innovation
SAP	Systemanalyse Programmentwicklung
SaaS	Software-as-a-Service
SME	Small and Medium-Sized Enterprise
SENAC	Serviço Nacional de Aprendizagem Comercial
SIC	Standard Industrial Classification
SLA	Service-level Agreement
S-D	Service-Dominant Logic
SDT	Self-Determination Theory
SPRU	Science Policy Research Unit
SQL	Structured Query Language
SSI	Sectoral Systems of Innovation

SWOT	Strengh, Weakness, Opportunity and Treat
TB	Terrabyte
TIM	Telecom Itália Mobile
TQM	Total Quality Management
TSI	Technological Systems of Innovation
U.S.	United States
USA	United States of America
VRIO	Value, Rarity, Imitability and Organisation
YARN	Yet Another Resource Negotiator
WEF	World Economic Forum

ABSTRACT

The University of Manchester
André A. Gomes De Souza
Doctor of Philosophy (PhD)
Analysing and Exploring Drifts in Innovation Streams within Open-Source
15th December 2020

Outbound OI is informed by the idea that organisations should use alternative pathways to externalise their knowledge and commercialise their technologies. Ambidexterity suggests a high level of balance between opposite, competing and contrasting objectives such as exploration and exploitation, as well as radical and incremental innovation. The study focuses on outbound OI processes of technology creators and consumers in their current markets and the strong mutual interactions between the literature of outbound OI and ambidexterity. Innovation streams discuss patterns of innovation and are one way to mobilise exploration and exploitation in service and product development.

The theory on which the thesis is developed views innovation as an evolutionary system. This vision is applied to the Apache Hadoop, the industry-standard ecosystem for the analysis of big data. Through an interpretive case study, the thesis investigates outbound OI processes in four case studies in six industry sectors in 25 organisations in Brazil.

The work contributes to the research on ambidexterity in outbound OI processes in technology creators and consumers in their current markets. It suggests an interrelationship between the code base of the Apache Hadoop distributions (community and enterprise) and the way firms may innovate (discontinuous, architectural and incremental). Additionally, the thesis adds additional case studies on outbound OI processes in digital service platforms.

DECLARATION

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENT

http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses.

ACKNOWLEDGEMENTS

My utmost gratitude goes to my academic supervisors. This work would not have been possible without the patient and supervision of Dr Khaleel Malik and Dr Dimitri Gagliardi. I have learned so much from them, and only through their support and encouragement, I was able to grow as an academic during this challenging time of my life. I was struck by bad health during most of the time of this PhD and loss a close family member. Dr Khaleel Malik and Dr Dimitri Gagliardi showed a lot of understanding. They were there for me when I needed it. I wish that I can be to my students what they have been for me. I also want to say thank you for the support and inputs of Prof. Philippe Laredo and Dr John Rigby. The thesis benefited a lot from their critical thinking and inputs.

I gratefully acknowledge the funding received towards my PhD from CAPES (Brazilian Ministry of Education). The grants provided some financial stability and peace of mind. I want to extend my thank you to Dr Maria da Graça Ferraz de Almeida Prado and Dr Scheila Manica for helping me out with the application. Additionally, a special thank you to Dr Isabelle Velasquez for formatting the thesis. I would also like to thank all firms involved in transcribing the interviews and proofreading the work. A big thank you to all interviewees; this PhD would not have been shaped without the contribution made by the Brazilian participants.

The process of writing a PhD can become very lonely. Good friends, fellow peers and partner are essential for PhD researchers. Therefore, I would also like to acknowledge the contribution made by these people. Big thank you to Adrian Liechti, André Ribeiro Bicalho, Elina Rocha, Dr José Ignacio Pineda, Kerstin Schriever, Patrick Mola, Wesley Silva Gomes dos Anjos, Will Oliveira and Zé Aruanã Kouyaté (Gaio).

Finally, I would like to extend my gratitude to my mother, Natalia de Souza Wyss. She supported my decision to start a PhD in advanced age like no other person and encourage me to continue. The sacrifice that she made for my education and her unconditional love will never be forgotten. Last but not least, a big thank you to Bruno Wyss.

To Tita

25.08.1938 - 24.12.2016

R.I.P.

CHAPTER 1. INTRODUCTION

1.1 RESEARCH BACKGROUND

Big data, or the managing of large data sets through a parallel information technology (IT) infrastructure, is one of the growing areas in research. This new technological innovation has raised a lot of interest in the academic community. A simple search query in Google Scholar results in over 1 million hits. The rise of big data has made it possible to track the ideas, visions and opinions of millions of individuals. The increase of social-media-related data in the early 2000s also gave rise to innovative means of researching social and cultural practices (Manovich, 2011). At the other end of the spectrum, Google and Amazon are only two examples of very successful institutions that are not only pushing big data technologies forward but also thriving economically. Other data-intensive organisations – Facebook and Netflix – have accrued billions of pieces of user data (Bughin, 2016). Academia and the private sector are overwhelmed by the possible access to this vast amount of information.

In light of the rise of this new economy, the World Economic Forum (WEF) in 2012 acknowledged the importance of data in today's economy and developed a ten-page document called Big Data, Big Impact. The significance of this document is that it recognised data, for the first time, as a new type of financial asset – putting it side-by-side with other commodities such as gold or silver. Many firms have only just scratched the surface in big data technologies (El-Darwiche, 2014). This recognition alone demonstrates the significance of big data as a subject of theory and practice.

It is important to note that data is only one element of this new aspect of the economy. While data seems to be the main protagonist, there is the need to store, manage and process the vast quantity that is being generated. In the light of this, the Apache Hadoop, a collection of tools available from the Apache Software Foundation, has emerged as the standard technology for big data analysis. In its original and most uncomplicated set-up, it consists of the Hadoop Distributed File System (HDFS) used for storing, and MapReduce for processing. The Apache Hadoop first appeared in a research paper in 2004, MapReduce: Simplified Data Processing on Large Clusters (Dean and Ghemawat, 2004).

The management literature suggests that outbound OI (OI) can drive competitive advantage. Instead of depending solely on internal pathways to market, outbound OI suggests that firms search for external organisations with business models better suited to commercialise a particular technology (Chesbrough, 2002). According to MacCormack and Verganti (2003), firms creating new software operate in unpredictable and dynamic settings. They often employ an iterative method that stresses learning and adaptability to succeed. Thus, to take advantage of OI approaches, firms must have specific capabilities to aid in the process. This study provides insights into the outbound OI process and its linkages to ICT tools, which is an important but understudied issue. It explains where and how Apache Hadoop creators and consumers apply ambidextrous outbound OI capabilities in their current markets.

Outbound OI belongs to an established practice in studies on innovation. It consists of the flow of internal knowledge to external environments enabling other firms to take advantage of their knowledge (Enkel et al., 2009; Dahalander and Gann, 2010; Huizingh, 2011). Outbound OI concerns: (1) the selling or out-licensing of IP to other firms' markets; (2) creating spin-offs out of internal technologies in new markets; (3) commercialising internally developed technologies in the current markets (Chesbrough, 2003/2006; Enkel et al., 2009).

Scholars (Mortara and Minshall, 2011; Ziegler et al., 2013; Chesbrough and Bogers, 2014; Smith and Akram, 2017; Bogers et al., 2018, Lee and Kim, 2019; Salampasis and Mention, 2019) have highlighted the lack of research in outbound OI and expressed the need for complementary studies. Others (Helm et al., 2017) recommends to identify and analyse additional case studies. For Vanhaverbeke et al. (2014), there is little evidence of outbound OI in service digital ecosystems and little progress has been made (Chesbrough and Bogers, 2014). According to Hu et al. (2015, p. 47), "outbound OI […] remains a challenge for most firms", and it is not very clear yet why firms cannot improve their outbound OI capacities. The same can be said about the link between OI and ambidexterity. Organisational ambidexterity is defined as the ability of firms to explore new service and products with new skills and exploit current service and products with existing knowledge (Andriopoulos & Lewis, 2009). According to Hafkesbrink and Schroll (2014, p. 9), the connection between the two agendas has not been fully explored, and there is some evidence of "strong mutual interaction" between the two research streams. There is empirical and theoretical evidence showing that to open up the idea phase of the innovation processes, firms need to be more flexible and, thus, exploratory. However, for the phases of the innovation process which follow, firms need exploitative forms of organisational design.

Outbound OI has several dimensions (Randhawa et al., 2016) and is a multi-level discipline (Bogers et al., 2017), leaving significant gaps (West and Bogers, 2014). It adds different contexts and levels of analysis to the study design, necessitating further theory building (Bogers et al., 2017). Outbound OI is a naturally dynamic process. Therefore, the research must include dynamic aspects (Appleyard and Chesbrough, 2017). The outbound OI literature discusses different variables that may influence the relationship between outbound OI and performance at varying levels of analysis, e.g. internal and external environment (Greco et al., 2017), firm size (Greco et al., 2017) and interdependencies between organisations and various stakeholders in an innovation ecosystem setting (Bogers et al., 2017).

The last ten years have seen a shift in corporate thinking about R&D, with the notion of doing everything in-house becoming obsolete (Appleyard and Chesbrough, 2017; Berchicci, 2013; Bianchi et al., 2016; Gassmann, 2006; Salter et al., 2015). Innovation has become progressively collaborative (Faems et al., 2005; Mention, 2011). Outbound OI is linked to firms gaining a competitive advantage when transferring information to the outside world (Cassiman and Veugelers, 2006; Hung and Chou, 2013). Outbound OI studies explore how firms export technical know-how (Hung and Chou, 2013; Naqshbandi et al., 2015; Naqshbandi et al., 2016; Parida et al., 2012), intellectual property (Cassiman and Veugelers, 2006; Hung and Chou, 2013; Tsai and Liao, 2011; Van de Vrande et al., 2009) and knowledge to the external environment (Hung and Chou, 2013).

According Viveiro Lopes and Monteiros de Carvalho (2018), outbound OI, when compared to inbound OI is still less common and less studied, presenting a strong potential for future researchers (Hsieh et al., 2016). Outbound OI has become an essential component of many businesses' business models. The growth of the academic discipline can be attributed to a variety of factors, including the shortening of product life cycles, more global competitiveness and growing R&D expenses. Based on this premise, the research seeks to understand how innovation takes place within Apache Hadoop and how it evolves, taking into consideration the outbound model of OI.

Outbound OI entails using internal knowledge for external usage in existing or new markets, as well as forming spin-offs (Huizingh, 2011). By externalising their expertise and commercializing ideas and products, outbound OI allows businesses to gain a competitive advantage (Chesbrough, 2003; Lee and Yoo, 2019; Reed et al., 2012). Dynamic capabilities are the result of management routines and organisational processes that are developed and moulded via asset positions that allow the organisation to adapt to strategic change (Teece et al., 1997). Hence, ambidexterity and outbound OI are linked by dynamic capacities. Dynamic capabilities assist the organization in being ambidextrous and enabling the organisation's innovation processes to become more open (Lieshout et al., 2021). Apache Hadoop vendors are ambidextrous by design. While Apache Hadoop can explore the Apache Hadoop through their community distribution, they can exploit it through their enterprise versions. One possible way to mobilise 'ambidexterity' in products and service development is the lense of innovation streams (Tushman and Smit, 2002; Benner and Tushman, 2003; O'Reilly and Tushman, 2008; Tushman et al., 2010). According to Smith and Tushman (2005), innovation streams defines patterns of innovation. It builds on and expands products and service. It, in turn, makes innovation streams an appropriate tool to understand organisational change and innovation in these specific environments.

The work adds to current discussions on outbound OI and ambidexterity by expanding prior studies' empirical evidence of outbound OI processes of technology consumers and creators in the firms' current markets. Second, few empirical studies have been conducted to date on outbound OI processes in free and open-source software service digital platforms ecosystems. Third, although the literature on ambidexterity is extensive, a lack of studies on ambidexterity in outbound OI remains. Therefore, the study answers the recent call for more research on ambidexterity in outbound OI processes in-firm.

While the unit of analysis is the dynamic capabilities of technology creators and consumers to externalise their knowledge and commercialise their technology, the object of interest is the Apache Hadoop. The Apache Hadoop is a free and open-source software, and it fits in with (West and Lakhani, 2009; West and Gallagher, 2005) and extends the literature on OI (West and Lakhani, 2009) as OI belongs to free and open-source firms' internal strategies (West and O'Mahony, 2008). The Apache Hadoop has been chosen for this study because: (1) due to the high investment of different firms across industries, there is a lot of innovation taking place in the ecosystems. It is now possible to analyse relations and trends within businesses and sectors, unlike before. This, in turn, stimulates innovation and improves tools to market. Additionally, big data allows firms to track and forecast events and problems. Finally, big data has shown enormous potential in reducing costs. (2) The rise of a radically new technological innovation potentially changes the way interactions take place between the different actors; therefore, it is essential to understand these changes; (3) The ecosystem is the main protagonist of an international industry, which makes it an exciting tool for understanding local structures.

This study provides insights into the outbound OI process and its linkages to ICT tools, which is an important but understudied issue. It explains where and how the Apache Hadoop may help companies implement outbound OI processes. OI is a firm-centric approach to innovation that separates it from other studies of the user innovation literature (Baldwin and von Hippel, 2011; Bogers, Afuah, and Bastian, 2010; Bogers and West, 2012). Although every inbound OI attempt is associated with an outbound effort by another organisation (Huizingh, 2011; Tranekjer and Knudsen, 2012). Yet according to the scholars (Chesbrough and Brunswicker, 2013; Chesbrough and Crowther, 2006), not all actors need to be equally engaged. Outbound OI is considered under-researched in comparison to inbound OI (in Stanko et al., 2017).

1.2 RESEARCH OBJECTIVES AND QUESTIONS

OI is a commonly used concept in academia and business and has become a paradigm to systematise innovation (Bogers et al., 2018). OI consists of the combination of internal and external ideas, their commercialisation channels and the alignment of these new concepts with the firms' business models (Christensen et al., 2005; Lettl et al., 2006; West and Gallagher, 2006). There are two kinds of OI: inbound and outbound. While inbound OI stands for inward technology transfer and the integration of these ideas in the firms' business models (Chesbrough and Crowther, 2006) outbound OI is characterised by firms leveraging their ideas by allowing others to take advantage of the firms' knowledge (Enkel et al., 2009; Dahlander and Gann, 2010; Huizingh, 2011).

OI has evolved as a key concept in both academic research and corporate practice, and it is now gaining traction in the public policy realm as well. Effective policymaking based on open innovation must capitalise on the value-added by openness in science while also encouraging the investment required to convert open projects into new technologies and commercial models. Outbound OI necessitates companies allowing underutilised and unused ideas to leave the company and be used by others in their enterprises and business models. In comparison to inbound, outbound OI has received less attention and hence is less well understood, both in academic study and in industrial practice (Mortara and Minshall, 2011; Ziegler et al., 2013; Chesbrough and Bogers, 2014; Smith and Akram, 2017; Bogers et al., 2018, Lee and Kim, 2019; Salampasis and Mention, 2019).

Outbound OI is implemented to accelerate knowledge externalisation and technology commercialisation. It is critical to understand how firms strategise their outbound OI projects to match the incentives and actions of other actors in their ecosystems. Collaboration with other firms whose operations are intertwined is critical to enable innovation in highly linked ecosystems (Adner and Kapoor, 2010; Iansiti and Levien, 2004; Pisano and Teece, 2007). Vanhaverbeke and Cloodt (2014) argue that the essential relationship between a firm's strategy and its outbound OI operations merits further research. Although further studies have helped to illuminate the organisational implications of outbound OI operations and the problems they face (Chiaroni et al., 2011; Mortara and Minshall, 2011), there are still underexplored issues regarding their underlying processes. Existing research has concentrated mainly on the firm-level drivers of outbound OI, ignoring their crucial interdependence with the activities of other companies in the ecosystem and their project-related decisions (Vanhaverbeke et al., 2014). The need to add to firm-level outbound OI studies with research at other levels has been repeatedly stressed (West et al., 2006; Chesbrough and Bogers, 2014), but little progress has been made so far (Masucci et al., 2020).

Ambidexterity is the ability to maintain a high level of balance between opposite, competing and contrasting objectives such as exploration and exploitation, as well as radical and incremental innovation (March, 1991; Levinthal and March, 1993; Tushman and O’Reilly, 1996, Simsek, 2009). Vendors commercialise the Apache Hadoop in two ways: through community and enterprise distribution. While vendors explore the technology through community distribution, they exploit it through the enterprise version. Consequently, Apache Hadoop vendors are ambidextrous by design. The two research questions investigate the conditions under which outbound OI and ambidexterity (dynamic capabilities) can be mobilised to understand how Apache Hadoop creators and consumers externalise their knowledge and commercialise their technologies in the firms’ current markets. Based on innovation streams (Tushman & Smith, 2002; Benner & Tushman, 2003; O’Reilly & Tushman, 2008; Tushman et al., 2010), this work examines how Apache Hadoop creators and consumers apply outbound OI to exploit and explore their knowledge and technologies, and how these dynamic capabilities influence the evolution of the ecosystem over time. The objective of this work is to illustrate the transition from exploitation to exploration and vice versa (from discontinuity to a dominant design), and how these dynamic capabilities are critical for the evolutionary system of the Apache Hadoop.

The study offers empirical evidence of the strong mutual interactions between the literature of outbound OI and ambidexterity in digital service ecosystems. The work pursued two sets of research objectives. First, it aimed to identify the contexts in which Apache Hadoop creators and consumers generate innovation streams in their current markets and synthesise the outbound OI processes of these firms. The research question associated with these objectives was:

How do Apache Hadoop creators and consumers generate innovation streams in their current markets?

Second, the study sought to unravel the underlying conditions under which the Apache Hadoop evolves from being a community base to an enterprise platform and synthesise the outbound OI processes of Apache Hadoop creators and consumers in their current markets. The research question associated with this objective was:

How can innovation streams evolve from community base to enterprise platforms?

1.3 RESEARCH METHODOLOGY

A qualitative research design was used to investigate how innovation streams are generated and evolve in firms. In line with Yin's (2013) ideas that qualitative research is more suitable for "how" questions in an under-explored field, the study follows bottom-up theory-building processes. The method used in the work answers to the requirements for extending theory-building in the outbound OI and the literature on ambidexterity in service digital ecosystems. 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. The advancement of new theoretical perspectives is fundamental in organisational research, and academics have often built theories by merging primary and secondary data. Glasser and Strass (1967) argue that the interrelationship between empirical reality and secondary data is a suitable laboratory for testing the validity of new theories (in Eisenhardt, 1989).

According to Eisenhardt and Graebner (2007), genuine empirical research starts with a review of related literature. This study relies on the authors' concept, identifies and suggests the research questions that aim to address the gaps. First, it frames innovation and focuses on evolutionary systems of innovation. Second, it covers inbound and outbound OI processes – the literature of outbound OI informs the research gaps. The third (free and open-source software) and fourth (service innovation) stages are complementary, and the foundations of understanding the demonstrator at investigation (the Apache Hadoop). Finally, the theoretical framework is organised from these four streams of literature.

The method used in this work is narrative inquiry. Case histories are used to understand how innovation in the Apache Hadoop ecosystems takes place and how the ecosystem evolves. The primary tool used in the investigation was interviewing, and the data was analysed based on theoretical sampling. Semi-structured interviews were the primary vehicle to direct and script the narratives. The semi-structured interviews were critical in helping Apache Hadoop professionals to think about and share their practices as they happened while working on various Apache Hadoop projects. The histories told by the participants were fundamental to structure the investigation and to collect personal accounts of the Apache Hadoop professionals in Brazil.

In order to do so, the researcher visited Apache Hadoop workshops in London, Berlin and São Paulo, in addition to studying industry use case publications and big data magazines. The workshops in São Paulo were of significant importance to networking with Brazilian professionals. The research approach is also in line with Rosen (1991). According to the author, interpretivism is a philosophy that believes the process of understanding lies in emerging in the world of those generating it (in Orlikowski and Baroudi, 1991).

The study also devoted effort to the real-life context in the case studies section, and by the end of the fieldwork, 25 semi-structured interviews had been conducted with Brazilian Apache Hadoop professionals from four different states in 25 different firms in six industry sectors. The participants were asked three sets of four questions about how innovation takes place within the Apache Hadoop ecosystem and how the ecosystem evolves from a community base to enterprise form of distribution. The interviews were conducted in the country's official language (Brazilian Portuguese), audio-recorded and later transcribed. The emerging data were later cross-checked and triangulated against available data about the Apache Hadoop ecosystem.

1.4 INTENTED CONTRIBUTIONS

The work intends to make contributions to some theoretical and practical managerial aspects of research in this field.

1.4.1 Theoretical Contributions

The work aims to contribute to two areas of research: outbound OI and the literature on ambidexterity within digital service ecosystems, in particular free and open-source software. Based on my case histories, informed by the research questions, the work finds empirical confirmation to extend the theory on outbound OI and ambidexterity three-fold:

(1) The study mobilises the role played by the different classes of innovation in the evolutionary process within service digital ecosystems creators and consumers in their current markets (from discontinuity to the development of a dominant design). The study adds to the literature of outbound OI with empirical evidence of an interrelationship between the operability and flexibility of the code base of the two different distributions of the Apache Hadoop ecosystem (community and enterprise) and the way firms may innovate (discontinuous, architectural and incremental). Additionally, it answers to recent calls for additional examples of outbound OI processes in different industries.

(2) The study organises processes of knowledge externalisation and technology commercialisation of creators and consumers of service digital ecosystems in their current markets. The study sheds lights on the interdependencies with the activities of firms in service digital ecosystems.

(3) The study articulates the essences and environments under which ambidextrous dynamic capabilities in outbound OI occur within service digital ecosystems. The study contributes to the literature of outbound OI and ambidexterity with empirical evidence of the strong mutual interactions between the two streams of literature. It also addresses recent calls for additional studies on ambidexterity in outbound OI.

1.4.2 Practical Implications

Implications are drawn for the broader domain of outbound OI and ambidexterity with particular reference to its applications to service digital ecosystems. More specifically, two lessons are drawn to the field. First, the thesis concludes that the community distribution offers more flexibility; therefore, it is more suitable for discontinuous and architectural innovation. The enterprise distribution offers more operability due to its technological maturity, therefore, more ideal for incremental innovation. The work offers recommendations to managers and firms' owners to define the types of innovation they want to pursue before they align their innovation with internal processes and technologies.

The second managerial implication is addressed explicitly to digital service ecosystems with a diversity of programming languages used. It is critical to understand that the right computing language is not only crucial to the development of the right workflow but also for accessing the right human resources. By committing to the right computing language for the digital service ecosystem in use, firms can broader their scope for specialised human resources.

1.5 THESIS STRUCTURE

The thesis comprises eight chapters, including the introduction. Chapter 2 outlines what innovation is and how it is defined in the innovation management literature, with a focus on its evolutionary perspectives. It moves on to discuss the OI literature, which is separated into inbound and outbound. Based on the literature on outbound OI and organisation ambidexterity (the ability of firms to mobilise their dynamic capabilities to gain competitive advantage), research gaps are identified, in combination with the two research goals presented in this chapter to inform the research questions. Chapter 2 also reviews the literature of free and open-source software and service innovation that is of significant relevance for understanding the demonstrator of the study, introduced in Chapter 5. Chapter 3 informs the theoretical framework of the study presented in detail. Chapter 4 describes the methodology and research design of the thesis. Chapter 5 introduces the Apache Hadoop ecosystem and opens with an explanation of what it is the Apache Hadoop? It continues with the importance of free and open-source software for big data, presents an overview of the Apache Hadoop architecture and ends with an overview of its vendors and key actors.

Chapter 6 presents the narrative of the four case histories by examining technical and business cycles and the outbound OI processes of Apache Hadoop creators and consumers in their current markets. Chapter 7 summarises the main findings related to the research questions. Finally, Chapter 8 closes this thesis by discussing the theoretical and managerial implications of the study, presenting some limitations and future research avenues.

CHAPTER 2. LITERATURE REVIEW

2.1 INTRODUCTION

This chapter explores the literature which has informed the thesis. The section commences by introducing what innovation is and how it is defined by different scholars in the discipline of innovation management. The first section also introduces the typologies of innovation and evolving models of innovation. The second section frames the literature discussed is OI and the third focuses on outbound OI. Here is where the research gap is identified. The fourth frame of literature is free and open-source software. The literature of free and open-source software is critical for the understanding of the demonstrator of the thesis, the Apache Hadoop, introduced later in this work in Chapter 5. The final part focuses on service science, which covers topics such as what is service science; the goods and service debate; service-dominant logic (S-D logic); service innovation; service ecosystems; and ends with ambidexterity in service innovation. The literature on ambidexterity is present throughout the chapter, and it is meant to be the linking element to the theoretical framework of the thesis, innovation streams, discussed in Chapter 3.

This chapter does not discuss gaps in the literature in the individual sections, though the critiques are presented in Section 2.4.7 Research Gap in Outbound OI. The literature review is used to inform the theoretical framework.

2.2 WHAT IS INNOVATION

What is "innovation"? This is a valid question and inspires vibrant dialogues and a wide variety of definitions. Due to the outstanding achievements of various scholars in understanding what "innovation" means, how innovation takes place and diffuses, we now have a much-improved subject competency.

Innovation is nothing new, and probably as old as humankind. There is something possibly inherent in us that makes us think up new and better ways of doing things, and putting these ideas into practice. Without this desire to improve we would probably be living in a vastly different world. Nevertheless, innovation has not always enjoyed the intellectual attention it deserves. Interest in the role of innovation in economic and social change has increased in recent decades, particularly within the social sciences, and has cross-disciplinary implications. This has resulted in a very diversified innovation literature (Fagerberg, 2006). In the traditional Schumpeterian definition, technical change is characterised as "a historic and irreversible change in the method of production of things" and "creative destruction" (Schumpeter, 1934). Table 1 summarises a possible implementation according to the classical interpretation (Kotsemir et al., 2013).

Table 1.A Possible Implementation According to the Traditional Schumpeterian Interpretation

A Possible Implementation According to the Traditional Schumpeterian Interpretation
goods that are new to consumers or of higher quality
methods of production that are new to specific industries and economic activities
opening of new markets; new sources of raw materials; new forms of competition

Source: Kotsemir et al., 2013.

Historians have frequently used “innovation” to explain different phases of economic development, and associated innovation with one of the main factors for the transitions between different epochs. Innovation seems to be one of the main drivers of a nation’s economic development. It is responsible for the advancement of cost-reducing processes, the launch of new products and services, and most importantly, for the creation of new ways of organising the activities of firms.

The Industrial Revolution in Britain in the late eighteenth century is one of the most significant examples of this: the shift from manual labour to manufacturing (Hudson 2004). Innovation was noticeably the key factor for the “new industrial” age, with steam engines and railways. The steam engine led to better iron and steel manufacture, which in turn enabled innovations in engineering, bringing a significant stimulus to the European economies (Ville, 1990). Significant advancements in scientifically-based industries characterised the second industrial revolution in the late nineteenth century (Pierenkemper and Tilly, 2004). Later in the twentieth century, American firms extended German technologies into organisational and marketing innovations. Cars were mass-produced on assembly lines and sold by third party dealers under a multi-divisional organisational structure (Chandler 1966). According to Fruin

(1992), by the 1950s Japanese firms started competing with Europe and North America due to their innovative manufacturing systems: lean production, new approaches to labour management, and the development of just-in-time contracting (Ville, 2011).

2.2.1 Definition of Innovation

As seen in the previous section, innovation has played an essential role in shaping the world we live in today. Even before the word 'innovation' was used in the way it is understood today, processes that are closely linked to innovation, and economic and technological change, were already recognised as critical (Lorenzi et al., 1912; Veblen, 1899; Schumpeter, 1934). The management of innovation is perceived as crucial and of interest for practice and theory across industries and the management disciplines. Baregheh et al. (2009) have pinpointed that different disciplines view innovation very differently, and they recommend distinct definitions. Table 2 illustrates some examples of definitions used in the discipline of organisational innovation (Baregheh et al., 2009).

"Innovation" can be understood in two ways: (1) the introduction of anything new, or (2) a novel concept, technique, or gadget" (Merriam-Webster, 2017). However, significant discrepancies exist between these two meanings. The first definition focuses on innovation as a result and the second as a process (Kahn, 2018).

Table 2. Examples of Definitions of Organisational Innovation

Authors	Definition
Thompson (1965, p. 2)	"Innovation is the generation, acceptance and implementation of new ideas, processes products or services".
West and Anderson (1996) Wong et al. (2008, p. 2)	"Innovation can be defined as the effective application of processes and products new to the organization and designed to benefit it and its stakeholders"
Kimberly (1981, p. 108)	"There are three stages of innovation: innovation as a process, innovation as a discrete item including, products, programs or services; and innovation as an attribute of organizations."
Van du Ven et al. (1986)	"As long as the idea is perceived as new to the people involved, it is an 'innovation' even though it may appear to others to be an 'imitation' of something that exists elsewhere".
Damanpour (1996, p. 694)	"Innovation is conceived as a means of changing an organization, either as a response to changes in the external environment or as a pre-emptive action to influence the environment. Hence, innovation is here broadly defined to encompass a range of types, including new product or service, new process technology, new organization structure or administrative systems, or new plans or program pertaining to organization members."
Plessis (2007, p. 21)	Innovation as the creation of new knowledge and ideas to facilitate new business outcomes, aimed at improving internal business processes and structures and to create market driven products and services. Innovation encompasses both radical and incremental innovation.

Source: Baregheh et al., 2009.

Since 2005, the Oslo Manual's articles 146 and 150 have been used to define innovation in the business sector for statistics purposes. Product and process innovation must be 'new or significantly enhanced,' while the two approaches must be 'new.' (Gault, 2018). Table 3 extends table two with two additional by definitions.

Table 3. Extensions of Definitions of Organisational Innovation

Definition
A product has to be 'introduced on the market' and a process or method has to be 'brought into actual use in the firm's operation'. The innovation takes place the moment the two conditions have been met.

Source: (Gault, 2018).

Although there are many ways of thinking about and defining innovation, a significant differentiation is usually made between invention and innovation. While invention seems to be the first manifestation of an idea for a new product or process, innovation is the first endeavour to put it in practice. Often the two words are used interchangeably, and it is sometimes difficult to separate one from the other. But there is a substantial time lag between the two, quite often even a decade (Rogers, 1995). This is because of the path an idea takes to its implementation. Another critical difference is that inventions can take place anywhere, in universities or a garage, while innovations often happen in firms by combining several types of knowledge, capabilities, skills, and resources. Firms are in a better position to innovate because they have the production knowledge, skills and facilities, market knowledge, and distribution channels. Most importantly, they have the capital to take the risks (Farberg, 2006). There are many different models, frameworks, classifications and definitions of types of innovations. Therefore, it is extremely difficult to understand the different types and their definitions, how they are used by various academics and the relationships between them. Table 4 summarises the definition of innovation and invention.

Table 4. The Definition of Innovation and Invention

Authors	Innovation	Invention
Freeman , 1982	Innovation is the introduction of change; something new.	Invention is the creation of a new device or process.
Senge, 1990	An idea becomes an innovation only when it can be replicated on a meaningful scale at practical costs	An idea has been 'invented' when it is proven to work in the laboratory.
Rouse, 1992	Innovation is the introduction of change via something new.	Invention is the creation of a new device or process.
O'Sullivan and Dooley, 2009	Innovation is more than the creation of something novel. Innovation also includes the exploitation for benefit by adding value for customers.	An invention need not fulfil any useful customer need and need not include the exploitation of the concept in the marketplace. Invention is often measured as the ability to patent an idea.

Source: Rowley et al., 2011.

Another basic definition of innovation is "the process of converting an idea or invention into an item or service that provides value or consumers willing to pay for it." For an idea to be considered innovative, it must be repeatable at a low cost and meet a particular need. The purposeful application of knowledge, creativity, and initiative to extract increased or different values from resources is referred to as innovation. It includes all processes by which new ideas are formed and translated into meaningful goods or service. From a business perspective, innovation often occurs when a company's views are deployed better to satisfy the requirements and expectations of its clientele (Mamasioulas et al., 2020). Table 5 extends the definition of innovation and invention based on Varadarajan (2018) and Caldas Vianna, (2021).

Table 5. The Extension of the Definitions of Innovation and Invention

Innovation	**Invention**
Innovation is the successful commercialization of an idea into a new product, process, or practice (Varadarajan, 2018).	Invention is a new arrangement of objects or other concepts used to describe an occurrence or suggest a new way of creating things (Caldas Vianna, 2021).

Sources: . Varadarajan, 2018; Caldas Vianna, 2021.

2.2.2 Typologies of Innovation

The importance of innovation for organisations seems to be acknowledged across industries and disciplines. Organisations need to innovate to adapt to changing customers' demands and lifestyles, and to take advantage of the opportunities that arise from new technology and changing marketplaces, structures and dynamics. Schumpeter (1950) already argued that organisations need to innovate, and a key concept in the literature of innovation is "type of innovation" Rowley et al. (2011).

Both external and internal limitations continually impact a firm's ongoing innovation decisions, which, in turn, further constrain or empower the firm's future options (de Jong and Marsili, 2006; Evangelista, 2000; Klevorick, Levin, Nelson, and Winter, 1995; Pavitt, 1984). Thus, firm-level innovation is influenced to some extent by the existing structure, form, and dynamics of the surrounding industrial environment and, in part, by the organisation's skills and decisions (Wozniak, 1993). Over time, both external factors and internal constraints will change. This often happens because of noise and random events and because of people's choices. Any scientific pursuit involves an analysis of typologies, which attempt to condense the complexity of the researched phenomenon into a few fundamental categories. According to de Jong and Marsili (2006), it permits the occurrence to be explained and anticipated based on the existence of a predefined set of systematic indicators for each category created (Gaskin et al., 2017).

The basic interpretation of innovation (shapes and typology) was established by the Organization for Economic Cooperation and Development (OECD) in a series of manuals. According to the OECD (2005, p.46) innovation means: "the implementation of a new or significantly improved product (goods or services) or process, a new marketing method, or a new organisational method in business practices, workplace organisation or external relations" (Kotsemir et al., 2013). There are many different models, frameworks, classifications and definitions of types of innovations. Therefore, it is extremely difficult to understand the different types and their definitions, how they are used by various academics, and the relationships between them. Table 5 seeks to summarise earlier models, classifications, typologies and frameworks of "types of innovation" based on Rowley et al. (2011).

In every industry, external variables interact with internal factors, resulting in businesses searching for and locating themselves inside niches within an industrial ecosystem (Miles et al., 1978). These variables include market volatility, expectations, and pressure, the firm's size, critical relationships within the organisation, path dependency (Cohen and Levinthal, 1990; Lyytinen and Damsgaard, 2010), mutual dependencies between critical strategic factors and thrusts (Miller, 1986), and mimetic and normative pressures (DiMaggio and Powell, 1983). In general, these elements contribute to developing an organisation's "selection environment" (Nelson and Winter, 1977). Additionally, other sets of selection restrictions emerge from internal variables such as imprinting (Marquis and Tilcsik, 2013) and cognitive inertia, which affects the firm's fundamental attitude toward innovation (Gaskin et al., 2017). Table 6 summarises the different types of innovation.

Table 6. The Different Types of Innovation

Perspective	Models	Literature
The models is divided in organisational structure, production process, people, and product/service.	Early Models	Knight, 1967
Binary models were proposed in the 1970s and 1980s. They discuss administrative, technical, incremental, radical, product, and process, as types of innovation.	Binary Models	Bantel and Jackson, 1989; Daft, 1978; Damanpour, 1991; Damanpour and Evan, 1984; Evan,1966.
More recently a number of integrative models have been proposed all of which identify a number of different types of innovation.	Integrative Models	Oke et al., 2007
The framework discusses position, process, product, and paradigm innovation. The framework aims at building a coherent knowledge base around the concept of innovation, and the practice and execution of	Recent Models	Francis and Bessant, 2005

innovation in organizations.		
In addition to product, process, and service innovation, it includes the following types of innovation: (1) new venture division, a new internal communication system and new accounting procedure; (2) Management Organizational Innovation such as TQM (total quality management), BPR (business process re-engineering); Quality Circles, just-in-time (JIT) (3) commercial/marketing innovation such as Direct Marketing.	Recent Models (Engineering)	Trott, 2005

Source: Rowley et al., 2011.

2.2.3. *Evolving Models of Innovation*

Theories on processes of innovation are organised as linear or systems oriented. There are critical similarities and linkages to consider when discussing the two classifications. On the one hand, linear views of the innovation process endorse a supply-side inclination in innovation policies. On the other hand, systemic views of innovation allow a much more effective view on the demand side, in term of theoretical and policy relevance. The linear models of innovation dominated the debate on processes of innovation throughout much of the period since World War II (Kline and Rosenberg, 1986). A linear approach means "science leads to technology and technology satisfies market needs" (Gibbons et al., 1994). It assumes that commercial research and development, as an applied science, follows a uni-directional process from ground scientific research to commercial applications. There is no dialogue with the subsequent stages of the innovation process such as product development, production, and marketing. On the other hand, systems-oriented views on innovation acknowledge the complex reliance on and the potential for various interactions between the different aspects of the innovation process (Edquist and Hommen, 1999).

According to Edquist (1997), the systems of innovation approach is critical because it enables us to explain, comprehend, clarify and shape the process of innovation. Most importantly, it makes it possible to determine the factors that shape and influence innovation. The system of innovation approach is built upon the theories of interactive learning (Lundvall, 1992) combined with approaches of the evolutionary theory of technical change (Saviotti and Metcalfe, 1991; Nelson and Winter, 1982). According to Chang and Chen (2004) there are three main approaches: (1) the national approach (Freeman, 1987; Lundvall 1992; Nelson, 1993); (2) the technological/sectoral approaches (Carlsson and Stankiewicz, 1991; Breschi and Malerba, 1997); (3) the local/regional approaches (Cooke et al., 1997; Braczyk et al., 1998; De la Mothe and Paquet, 1996).

System innovation's primary application areas and contributions are drawn from two aspects. The first two components address regional development and science, technology, and innovation policy (Edquist, 2001; Boschma, 2004). The other two components draw on the system innovation study's evaluation of knowledge flows (OECD, 1997). Thus, system innovation research contributes significantly to the knowledge management literature, particularly identifying and assessing the value of new categories of information and their associated flows inside a system innovation (Bassis and Armellini, 2018).

2.2.3.1. National Systems of Innovation

The term ‘systems of innovations’ (SI) was coined by Freeman (1987) in his studies on national systems of innovation (NSI) (Freeman, 1987/1995; Lundvall, 1992; Braczyk and Heidenreich 1998; Nelson, 1993). According to Freeman, NSI is: “the network of institutions in the public and private sectors whose activities and interactions imitate, import, modify and diffuse new technologies” (Freeman, 1987, p. 1). Lundvall extended the concept and described NSI narrowly as: “organisations and institutions involved in searching and exploring—such as research and development (R&D) departments, technological institutes, and universities” (Lundvall, 1992, p.12), and broadly as: “a system of innovation ... constituted by elements and relationships which interact in the production, diffusion and use of new and economically useful knowledge” (Lundvall, 1992, p. 13). According to Nelson and Rosenberg (1993, pp. 2, 3), NSI means: “a set of institutions whose interactions determine the innovative performance” (Chang and Chen, 2004).

Schumpeter introduced the idea of "innovation" in *The Theory of Economic Development* at the turn of the twentieth century. He described and pioneered the theory of innovation. According to him, innovation is the formation of new production functions, a novel combination of production methods. Schumpeter's revolutionary approach encompasses technological innovations and those that increase the efficiency of component configuration. In the 1980s, Lundvall introduced the notion of the "innovation system", and academic circles have since conducted a systematic and comprehensive study of the innovation system's formation. Due to the variety of a country's regions, disparities in innovation performance might exist not just across nations but also between regions. Cooke (1992) pioneered the notion of a regional innovation system (Li and Zhang, 2020).

2.2.3.2. Technological or Sectoral Systems of Innovation

Technological systems of innovation (TSI) or sectoral systems of innovation (SSI) investigate the knowledge associations between firms and their results uniquely from technical linkages. Both systems, SSI and TSI, underlie the economic dynamics of technology development (Zuscovitch, 1986) and the significance of inter-industry technology flows (Carlsson and Stankiewicz, 1991; Breschi and Malerba, 1997; Pavitt, 1984; Porter 1990). Carlsson and Stankiewicz (1991, p. 111) characterise technological systems as: “networks of agents interacting in a specific technology area under a particular institutional infrastructure to create, diffuse and utilize technology focus on knowledge, information and competence flow”. Breschi and Malerba (1997, p. 131) define SSI as: “the specific clusters of the firms, technologies, and industries involved in the generation and diffusion of new technologies and in the knowledge flows that take place amongst them”. The interdependence within industries is the result of the accumulation of generated knowledge and the interconnections among technologies and industries (Maskell et al., 1998). SSI challenges the view that all technological or sectoral systems are homogeneous. Additionally, SSI stands for a clear understanding of the nature of technology (tacit or codified) and the relation between science and technology (Metcalfe, 1995). According to Archibugi et al. (1999), sectors and techniques have their dynamics and are of unmeasurable importance. There are differences in technical change between manufacturing sectors. This is due to the variation in sources of technology, the involvement of user needs, and means of appropriate benefit (Pavitt, 1984). SSI believes that some industries are defined by a few organisations placed in particular geographical localities in which they can work together on specific aspects of the innovation process but compete with other regions within and across countries. SSI investigates the inter-industry, interdependent relationship not only at local and national levels but also in broader global systems (Chang and Chen, 2004).

Numerous efforts have been made to quantify innovation, for instance, employing expenditure in R&D and the acquisition of patents as innovation indicators. However, this method uses the business as the unit of study. It focuses only on inputs and outputs, excluding consideration of the innovation process among sector participants. Under the sectorial system of innovation approach, the innovative performance of businesses cannot be evaluated based only on their efforts and accomplishments individually (Malerba and Orsenigo, 1996). Rather, innovation occurs due to the interaction of people with similar or dissimilar institutional backgrounds (Malerba, 2006). The early pioneers of the innovation systems (method conducted analyses at the national level, including R&D activities and universities, research institutions, government agencies, and policymakers (Freeman, 1987; Lundvall, 1992; Nelson, 1993). Lundvall defined innovation systems in 1992 as a collection of separate components and their interactions. This concept recognises that economic, political, and cultural issues impact how innovation systems are organised since they affect innovation efforts' scope, direction, and success. Malerba and Orsenigo (1996) developed the SSI approach to comprehend the unique characteristics of an industry better. According to Malerba (2006), a sector is a collection of activities linked by a shared set of goods aimed at a particular or rising demand and with a similar knowledge base. Thus, the SSI approach views innovation as a process involving systematic interactions amongst diverse players to create and share valuable information about innovation (de Santos e Silva et al., 2019).

2.2.3.3. Regional or Local Systems of Innovation

Since the 1980s, a pattern of localised production systems has evolved, which has had a growing influence and relevance on economic geography and regional growth. New definitions, such as technological districts, innovative milieu, learning regions, and regional systems of innovation incorporate established theoretical and empirical constructs for those associated with industrial and regional growth. Based on conclusions made and lessons learned from particular case studies and regions over the last decades, some scholars (Maillat, 1998; Markusen, 1996; Doloreux, 1999) concentrated on the basic types of industrial systems and dealt with the processes and circumstances that make the regions successful or not. Considering that localised production systems serve various roles, the definition of regional systems of innovation (RSI) has been discussed in the literature as to how they may vary from other territorial systems. Its features have been critically discussed concerning their origin and characteristics (Edquist, 1997), system components (Lundvall, 1992), system performance and evaluation (Autio et al., 2000), and the conditions of its use as a framework to assist innovative regional policies (Cooke, 2000) (in Doloreux, 2002).

The term 'regional and local systems of innovation' (RSI) was coined in the late 1990s, coinciding with research in industrial economics, regional economics, and economic geography. The definition of RSI is explicitly or implicitly very diverse among these disciplines. Still, historically the concept of RSI has its origins in Marshall's industrial district (Marshall, 1932), Perroux's economic spaces (Perroux, 1950), Dahmén's development blocks (Dahmén, 1988), Camagni's innovative milieu (Camagni, 1991), and regional innovation systems (Cooke et al., 1997; De la Mothe, 1996). RSI can be associated with the response to the anticipated concerns with the local supply managerial and technical skills, accumulated tacit knowledge, and knowledge spill-over.

According to Saxenian (1991), Silicon Valley in California is a classic example of RSI. Her study illustrates inter-firm networks and the way these firms can share the cost and risks of developing new technologies, and how to promote mutual innovation among the firms involved. Theis inter-firm collaboration functions as a platform for information exchange in technology transfer. It also contributes to the mobility of human resources and development and networking within the region. The inter-firm network provides a successful and dynamic relationship with technological innovation (Chang and Chen, 2004).

The concept of regional innovation systems (RIS) has grown in popularity among academics and across diverse disciplinary domains, with worldwide political ramifications. According to Chung (2002), RIS is a valuable tool for efficiently and successfully administering national innovation systems since it can foster different sector-based innovation systems in other areas. Furthermore, RIS (Asheim et al., 2019; Asheim and Gertler, 2005; Cooke, 1992) include dynamic strategic alliances between private and public players, including decision-makers in political institutions, businesses, and other organisations (Andersson, 2013). According to Rondé and Hussler (2005), the RIS literature views innovation as an evolutionary and social collaborative learning process (Fernandes et al., 2020).

2.3. OPEN INNOVATION (OI)

2.3.1. Open Innovation Research Tradition

Alan Pearson and Derek Ball laid the ground for the R&D management field and were the pioneers in the subject area (Pearson et al., 1979; Griffiths and Pearson, 1973). Rothwell and Zegveld (1985) developed the network model of innovation. Carter and Williams (1959), Allen and Cohen (1969) and SPRU's Project SAPPHO (Rothwell, 1974) uncovered in different studies that a fundamental feature of innovative firms is the quality of the inward information. Other studies have revealed that organisations running their own R&D departments are better positioned to approach external knowledge (Tilton, 1971; Allen, 1977, Mowery, 1983; Cohen and Levinthal, 1989). Dahlander and Gann (2010) pointed to various examples of absorptive capacity (Cohen and Levinthal, 1990), complementary assets (Teece, 1986), and the exploration versus exploitation debate (March, 1991). Ziegler et al. (2013) coined the ideas of the origins and dynamics of production networks in Silicon Valley of descriptive capacity to complement the well establish concept of absorptive capacity and to characterise the firms' competence to externally exploit knowledge. Von Hippel's lead users (1986), Katz and Allen (1982) not invented here (NIH) syndrome, Mowery (2009) argument that closed innovation is rather the exception and Allen's (1983) discussion of the iron production industry in the 19th century in England were only a few examples of how innovation has rarely emerged from closed systems. Many scholars in the disciple of R&D management and innovation management reason that OI embodies nothing more than the rewrapping and the recombination of ideas and discoveries from the past four decades. Thus, it is fairly reasonable to question if OI is "old wine in new bottles" (Trott and Hartmann, 2009).

Current open innovation research covers a wide range of topics and domains, including small and medium-sized enterprises (SMEs), new units of analysis, various high- and low-tech sectors, non-profit organisations, and public policy (Cassiman and Valentini, 2016; Faems et al., 2010; Laursen and Salter, 2006). Because innovation is a fundamentally complicated and dynamic social process, integrating theory and practice has enormous significance. A variety of causes have contributed to the emergence of open innovation as a concept as well as a research subject and community. A core idea is that knowledge is broadly dispersed in the market. Additionally, the increased mobility of workers, more capable universities, declining US hegemony, growing access of startup firms to venture capital (VC), and the rise of the Internet, social media, and supporting information and communication technologies are some key factors that have amplified the importance of open innovation 19 (Bogers et al., 2018).

2.3.2. The Notion of Open Innovation

The notion of OI was invented by Henry Chesbrough from Berkeley University and gained international attention through the publication of the book OI: The New Imperative for Creating and Profiting from Technology, which appeared in 2003. Chesbrough claimed that firms in the 21st century would spend heavily on internal R&D, allocating their best employees. This, in turn, would allow the firms to develop the most innovative products and services. Most importantly, there would be a need for ideas to be protected with intellectual property (IP) strategies. The returns on the investments would be reused to finance further innovation in the firms' R&D. By the end of the last century, different factors had forced firms to rethink their closed innovation processes. Chesbrough suggested that, through OI processes, firms can and should go to market to commercialise external as well as internal innovation by using external and internal pathways (Chesbrough, 2003). OI is used to characterise "both a set of practices for profiting from innovation and a cognitive model for creating, interpreting and researching those practices" (West et al., 2006, p. 286). West and Gallagher (2006) determined four generic OI strategies: (1) Pooled R&D – shared R&D (requires a shift in culture); (2) Spinouts – a way of escaping large firm bureaucracies; (3) Selling complements – accepting commodification or developing differentiated products based on commodities; (4) Donated complements – general purpose technologies are sold so users can develop differentiated products (e.g. user toolkits) (Elmquist, 2009).

Over the last half century, R&D tasks have shifted considerably. In particular, the last two decades have highlighted many adjustments in the way organisations handle their technological assets and their R&D departments, although keeping secrets and working within organisational boundaries are still dominant and relevant. Table 6 summarises the different topics covered in the OI literature (Fredberg et al., 2008).

While the present open innovation paradigm continues to dominate, the human resource (HR) component and the meaning and influence of businesses' innovation milieus are critical factors. The innovative environment in which a business operates is deeply intertwined with its HR management and strategies for recruiting and keeping talent (Meissner and Shmatko, 2017/2018). Bringing talent to firms for innovative purposes is often more manageable than retaining talent and encouraging employees to perform exceptionally well. The latter is difficult because organisations lack competencies, not just those directly relevant to innovative activities but also in management and legal matters (Cascio and Aguinis, 2005; Van de Vrande et al., 2009). Firms' need for innovative capabilities and corporate management's increased expectations for innovation are critical factors in everyday business activity. Therefore, it is vital to rethink how innovation and HR management are structured internally to foster an organisational environment that equips employees to address these issues (Del Guidice et al., 2018). Table 7 summarises themes of OI and Table 8 extends Fredberg et al. (2008) with Usnam et al, (2018).

Table 7. Themes of Open Innovation

The Notion of OI	Chesbrough (2003ab, 2004, 2006ab); Chiaromonte (2006); Gassmann and Reepmeyer (2005); Gaule (2006); Gruber and Henkel (2006); Motzek(2007); West and Gallagher (2006); West et al. (2006)
Business Models	Chesbrough (2003c); Chesbrough (2007); Chesbrough and Schwartz (2007); Van der Meer (2007)
Organisation Design and Boundaries of the Firm	Brown and Hagel (2006); Chesbrough (2003b); Dahlander and Wallin (2006); Dittrich and Duysters (2007); Fetterhoff and Voelkel (2006); Jacobides and Billinger (2006); Simard and West (2006); Tao and Magnotta (2007)
Leadership and Culture	Dodgson et al. (2006); Fleming and Waguespack (2007); Witzeman et al. (2006)
Tools and Technology	Dodgson et al., 2006; Enkel et al. (2005); Gassmann et al. (2006); Henkel (2006), Huston and Sakkab (2006; 2007); Piller and Walcher (2006); Tao and Magnotta (2006)
Intellectual Property, Patenting & Appropriation	Chesbrough (2003a); Henkel (2006); Hurmelinna et al. (2005)
Industrial Dynamics and Manufacturing	Bromley (2004); Christensen et al. (2005); Cooke (2005); Vanhaverbeke (2006)

Source: Fredberg et al., 2008.

Chesbrough (2003a) argues that intensified rivalry and shorter product life cycles have resulted in the reduction of organisational growth, even though organisations are more likely to spend much more money with R&D. OI presumes that organisations can and should consider external designs, as well as ideas developed internally, when developing new products and/or service (Chesbrough, 2006). Additionally, firms need to align these concepts, whether from internal or external sources, with the firm's existing business models (Chesbrough, 2003a/2006; Chesbrough and Rosenbloom, 2002; Zott et al., 2011).

Table 8. Extension of Themes of Open Innovation

Extension of Themes of Open Innovation
Adoption of Open Innovation
Benefits of Open Innovation
Challenges of Open Innovation
Role of Networking
Sectoral Patterns
Role of Policymaking

Source: Usnam et al, 2018.

OI consists of the combination of internal and external ideas and their commercialisation channels (Christensen et al., 2005; Lettl et al., 2006; West and Gallagher, 2006). Inbound OI denotes inward technology transfer and portrays the tradition of integrating external ideas in one's own business models, and assumes that modern organisations cannot rely solely on their own R&D activities (Chesbrough and Crowther, 2006). Outbound OI advocates outward technology transfer and refers to external pathways to commercialise an innovation (Chesbrough and Crowther, 2006). Outbound OI is discussed in detail in section 2.2.6. Outbound OI. Table 9 illustrates some of Differentiation for OI relative to prior theories of innovation.

Table 9. Differentiation for Open Innovation, relative to prior theories of innovation

Equal importance given to external knowledge, in comparison to internal knowledge
The centrality of the business model in converting R&D into commercial value
Type I & II Errors (in relation to the business model) in evaluating R&D projects
The purposive outbound flows of knowledge and technology
The abundant underlying knowledge landscape
The proactive and nuanced role of IP management
The rise of innovation intermediaries
New metrics for assessing innovation capability and performance

Source: Chesbrough and Crowther, 2006.

2.3.3. Organisation Design, Firms' Boundaries and Openness

There is an established frame of literature in organisational theory based on setting a firm's boundaries in a manner that prevents it from depending on its environment and allows it to protect critical tasks (Thompson, 1967; Pfeffer & Salancik, 1978 [2003]; Aldrich, 1979; Santos & Eisenhardt, 2005). Since Schumpeter (1947), different scholars (Barnard, 1938 [1968]; Chandler, 1962; Myers and Marquis, 1969) have discussed innovation as the result of internal R&D capabilities in combination with inventions, patenting and the administration of the firm's innovation portfolio (Allen, 1977; Cohen & Levinthal, 1990; Clark & Fujimoto, 1991; Dougherty & Heller, 1994; Fleming 2001; Dougherty & Dunne, 2011 in Lakhani et al., 2013).

Table 10. The Contrasting Principles of Closed and Open Innovation

Closed innovation principles	Open Innovation principles
We have the best professionals in the field working for us.	Not all "best" professionals work for us. There are good professionals inside and outside our firm.
To be able to benefit from internal R&D, our firm needs to discover it, develop it and sell ourselves.	External R&D can create value for our firm. Internal R&D is important to extract some portion of the value.
What we discover we put first on the market	We do not need to initiate research to benefit from it.
The firm that first takes a service/product to market wins the most.	It is more important to develop a business than be the first.
The firm that develops the best ideas in the field will conquer the market.	The firm that can combine internal and external ideas will conquer the market.
It is important to protect our IP. Competitors cannot take advantage of our ideas.	The firm can benefit if others use our IP. We should commercialise our IP and buy other firms' IP if it fits our business model.

Source: Chesbrough, 2003.

The idea suggested by the authors is that innovation happens behind closed doors. It is vital to protect and control the innovation and its processes from actors outside of the firm's boundaries; this concept is known as "closed innovation". Most recently, different elements have challenged the well-established notion of closed innovation, although a combination of factors has contributed to the paradigm change. It seems that two were of significant importance: (1) the growing mobility of highly experienced and skilled people; (2) the ever-increasing presence of private venture capital (Chesbrough, 2003b). Table 10 illustrates the contrasting principles of closed and OI.

The work on OI contradicts the intellectual tradition discussed previously. Instead, it proposes that firms can prosper from innovation originating from outside the firm's boundaries. One possible way of handling the pressure of highly competitive environments and short life cycles is to share the development of high-tech products/services with other firms. Firms are increasingly inclined to build technological alliances at the core of their technology strategies. Strategic technical alliances have driven the formation of inter-firm networks, in which firms participate jointly on the development of new products and technologies. March (1991) associated the exploration of new opportunities and ideas with research, experimentation, risk-taking and innovation. According to Granovetter (1973), new ideas and opportunities often come into being through communication with partners (alliances) in different lines of business as the firms acquire new knowledge. Firms following an explorative path generally build weak ties (Granovetter, 1973). According to Burt (1992), exploration is associated with opportunistic behaviour. Weak ties allow a firm to link with two or more networks of firms and possibly take advantage of all the resources from the networks. Accordingly, exploitation of current knowledge and capabilities is bound to refinement, selection, production and execution (March 1991). When firms build an exploitation strategy, they can enhance and expand fundamental knowledge of established technologies and products. Widening the numbers of strong ties will augment basic knowledge of established techniques and products (Dittrich and Duyester, 2007).

There is extensive literature discussing the advantages of open over closed innovation (Dahlander and Gann, 2010; West and Bogers, 2013). Chesbrough (2003, p. 23) defined openness as "a paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as firms look to advance their technology". Chesbrough's definition is the most commonly used and highlights that valuable ideas can begin and be commercialised from inside or outside the firm. Table 11 summarises examples of empirical studies on openness based on Dahlander and Gann (2010).

A critical part of the process of innovation is the pursuit of new ideas that may have the potential for commercialisation. With this in mind, firms invest resources (labour, time and money) seeking new innovative opportunities. As innovation is path-dependent, the larger the search, the greater the ability to develop, put innovative solutions into practice and recombine new and existing knowledge, hence the importance of networks, communities, and links with other firms. The process described above redefines the boundary between firms and their surroundings. The literature on openness and interaction in innovation studies follows studies of organisational behaviour intellectually (Laursen and Slater, 2006).

External knowledge sharing entails additional difficulties. Because core information is crucial to establishing and maintaining a firm's competitive advantage, it must be preserved (Katila et al., 2008). Sharing and exposing an excessive amount of such core information may have a detrimental influence on a business's innovation performance (Baughn et al., 1997; Dahlander and Gann, 2010) the overall performance of the firm (Frishammar et al., 2015), and future competitive advantage (Norman, 2002). Firms recognise the need of preventing information spillovers and imitation (Ethiraj et al., 2008) and exercise caution in defining what sort of knowledge is important and how this knowledge is communicated (Connell and Voola, 2013). Simultaneously, disclosing specific quantities of information to partners (deliberately or accidentally) might be helpful (Alexy et al., 2013; Alnuaimi and George, 2016; Henkel, 2006). Leakages may be beneficial if they aid in disseminating information that results in new commercial prospects and boosts the buzz around a product or technology (Alexy et al., 2013).

Table 11. Examples of Empirical Studies on Openness

Study	Context	Key results	Focus
Chesbrough and Crowther (2006)	Low-tech or mature industries	OI is also present in low-tech firms. Leveraging external resources as a complement rather than a substitute for internal R&D	inbound – acquiring
Christensen et al. (2006)	Consumer electronics	The use of OI depends on the position in the innovation system and stage of the technological regime	inbound – acquiring
Laursen and Salter (2006)	Firms within manufacturing	Search and performance (breadth and depth) relate to each other curvilinearly	inbound – sourcing
Fey and Birkinshaw (2005)	Firms with R&D activities	How governance mode for external R&D, openness and knowledge affect R&D performance	inbound – sourcing
Henkel (2006)	Embedded Linux	Selectively revealing free and open-source software	outbound – revealing
West (2003)	Proprietary platform vendors	Strategies trade-off by open-source firms between appropriation and appropriability	outbound – revealing

Source: Dahlander and Gann, 2010.

The studies advocate that networks and knowledge exchange between firms can play a critical role in improving performance. The same can be said about research on evolutionary economics, which considers openness to the external environment beneficial for firms' innovation processes. Scholars in the discipline of evolutionary economics stress the role of "searching" in assisting firms to develop new combinations of technologies and knowledge (Nelson and Winter, 1982). The portfolio equips firms with an array of technological paths to choose from (Metcalfe, 1994). According to Nelson and Winter (1982) and Levinthal and March (1993), the strategies used in searching are also affected by the variety of technological opportunities at hand and the pursuit of other firms (Laursen and Slater, 2006).

The concept of openness assumes that organisations on their own are not capable of innovating. Organisations have to interact with other actors to amass ideas and resources from outside their boundaries to stay ahead of the competition (Chesbrough, 2003a; Laursen and Salter, 2006). The discussion questioned the importance of openness in innovation. It highlighted the permeability of firms' boundaries where ideas, resources and individuals flow in and out of firms. An external partnership can leverage the investments in internal R&D through broadening the opportunities of previously disconnected silos of knowledge and capabilities (Fleming, 2001; Hargadon and Sutton, 1997; Schumpeter, 1942). Openness is characterised by various types of relationship with external actors and is closely linked to the boundaries of organisations (Dahlander and Gann, 2010).

Sharing information without the expectation of money is inevitable in a connected business environment where frequent interactions with outsiders make it impossible to keep knowledge safe (Macdonald, 1993). Tesla's open attitude to intellectual property, which allows for developing complementary technologies that enhance Tesla's products, is a well-known current example of strategic disclosure (Tietze, 2017). Previous research indicates that organisations might anticipate several advantages from disclosing (Henkel et al., 2014). From a marketing point of view, revealing can improve reputation (Henkel, 2006), goodwill (West and Gallagher, 2006), brand recognition (Dahlander and Magnusson, 2008), and can be used to reach more people (West, 2006). From a technical standpoint, organisations may profit from crowdsourcing expertise rather than solving problems internally or engaging a specialist provider (Afuah and Tucci, 2012; Piezunka and Dahlander, 2015) (Verreynne et al., 2020).

In order to innovate, organisations have to become more permeable to external knowledge (Cassiman and Valentini, 2015). The antiquated Schumpeterian ideal of the isolated entrepreneur flying solo has been overtaken by the model of different stakeholders operating collectively in repetitive procedures of trial and error to develop and commercialise new ideas successfully (Schumpeter, 1942 [87]; Rosenberg, 1982; von Hippel, 1988; Freeman and Soete, 1997; Tidd et al., 2000). The software industry has gone through a major shift. Free and open-source software is based on new types of business model that often rely on collective creativity through OI. According to Chesbrough (2007), if organisations are to make strategic decisions based on innovation communities, ecosystems, networks, and how the new business models may affect the firms' competitive advantage, firms will need a new approach to strategy—what Chesbrough (2007) calls "open strategy".

2.3.4. Open Innovation Business Models

Business models play a critical role in OI processes. According to Chesbrough (2007), a better business model will always beat a better idea or technology. In order to innovate it is important to understand what a business model is, to scrutinise the existing business model options and how to improve them. Business models fulfil two critical duties: (1) It describes an array of tasks—from logistics to consumer satisfaction. Table 12 summarises the functions and frameworks of business models.

Table 12. The Functions and Frameworks of Business Models

Articulate the value proposition: that's to say, the value generated by product for consumers
Identify a business segment; namely the consumers to whom the product is beneficial and for what reason
Establish the configuration of the supply chain required by the organisation to construct and deliver the bid, and define the external assets needed to sustain the firm's role in this chain. It involves the manufacturers and clients of the firm and can apply to the end user from the raw materials.
Specify the firm's revenue generating mechanism and evaluate the cost structure and benefit efficiency of the bid based on the value proposition and supply chain structure selected.
Describe the firm's role within the value network (ecosystem) that connects suppliers and consumers, including the detection of possible add-ons and competitors.
Formulate the strategic approach for the creative business to win and take advantage of its competitors.

Source: Chesbrough, 2007.

Business model innovation (BMI) is defined as a method (Schallmo and Brecht, 2010; Berglund and Sandström, 2013; Matzler et al., 2013; Foss and Saebi, 2017) for the creation of a new business model for a firm (Björkdahl and Holmén, 2013; Foss and Saebi, 2017) or an entire industry (Foss and Saebi, 2017) (Santos et al., 2009; Schallmo and Brecht, 2010; Foss and Saebi, 2017). BMI may also refer to the outcome of an innovation programme that replaces or revises an organisation's current business model (Mitchell and Coles, 2003; Foss and Saebi, 2017) or that entirely transforms the existing business model (Lindgardt et al., 2009). These changes may affect the value chain or the company's value offer to its customers or other partners (Wirtz, 2011; Matzleret al., 2013). As a result, conversations concerning BMI involve the degree of innovation (Amit and Zott, 2012; Lindgardt et al., 2009; Hacklin et al., 2018) (in Rachinger, 2018).

The tasks redefine the new product/service in such a way that they create value through the activities. (2) The business model acquires value from the activities described in (1) when firms develop and operate the tasks. There are different approaches to improving a business model but the best is by dividing the improvements into phases of business model advancements (Chesbrough, 2007). Table 10 summarises the functions of business models (Chesbrough, 2007). There is a frame of studies within the business strategy literature that tries to understand how firms compete effectively: Porter's five forces (Porter, 1980), generic strategy framework (Porter, 1980) and value-chain framework (Porter, 1985); generic building blocks framework (Hill & Jones, 2001); SWOT (Helms and Nixon, 2010); VRIO framework (Barney, 2002). The frameworks mentioned are only a few of the many different options that can help put the strategy into action (Richardson, 2005). The Business Model Framework (BMF) is a way to think about the improvements of a business model as a path from basic (not very valuable) to advanced (very valuable) models.

Applying BMF is one possible way to evaluate the current status of a business model and determine the next step forward (Chesbrough, 2007). Table 11 summarises the different Business Model Frameworks (Chesbrough, 2007). One critical instrument for innovating a firm's business model is through building co-development alliances. Furthermore, to keep the alliance alive, both firms need to specify very carefully their business objectives and align the business models of each firm accordingly. OI has proved to be beneficial to organisations that have adopted it, and the potential for business model innovation via co-development is significant. In contrast to business model innovation, a traditional business model is based upon the belief that product and service development is the result of internal technology

(R&D). The concept of sharing the research and/or development of new products or services can help significantly to reduce R&D expenses, widen innovation output, and grant access to new markets that may otherwise have been inaccessible (Chesbrough, 2003a).

BMI is critical for businesses since it must be effectively understood and implemented. It is vital to comprehend the fundamentals of BMI theory and the state of the art. It is imperative to notice developing trends in this sector to keep up with new ideas. This insight may benefit practitioners and scholars by indicating future business activities and study areas. Foss and Saebi (2017) sought to map the state of the art of BMI. However, their primary objective was to identify and correct gaps in BMI research, not to show its fundamentals. As a result, there is still no well-defined core body of research on BMI. While reviews summarise the literature, they do not examine developing patterns or identify key components (Filser et al., 2020).

Critical to the development of a co-development partnership strategy is the definition of the business objectives for partnering. The key elements of the co-development design may differ depending on the business objective. It is critical to align strategy, objectives and partnerships to secure results. According to Chesbrough and Schwartz (2007), every organisation has a network consisting of its supply chain, distribution system and customers. However, most firms disregard the importance of coherently detailing their own business model. Table 13 summarises the different types of business model frameworks.

Table 13. The Different Types of Business Model Frameworks (BMF)

Type 1:	Company has an undifferentiated business model.
Type 2:	Company has some differentiation in its business model.
Type 3:	Company develops a segmented business model.
Type 4:	Company has an externally aware business model.
Type 5:	Company integrates its innovation process with its business model.
Type 6:	Company's business model is an adaptive platform.

Source: Chesbrough, 2007.

One way to improve the co-development partnership is by having a clear vision of one's own business needs, and the degree of their alignment with one's business model (Chesbrough and Schwartz, 2007). Table 12 illustrates the different business objectives of co-development (Chesbrough and Schwartz, 2007). Chesbrough (2003) stressed the importance of the flexible use of several business models. According to the author, flexibility is critical to prevent the "Not Sold Here" syndrome. When an organisation can manage the development and adoption of additional business models efficiently, new profitable opportunities may arise. Business models can be viewed as a tool to transform technical aspects of a product or service into economic value. Business models focus on the fundamental aspect of what it takes to transform technology or specific know-how into (commercial) success. Business models act as a bridge between the technical domain—what do we deliver—with the social domain—how much value does our service/product provide our customers and how are we capitalising on it? This is of fundamental importance to business models in the OI paradigm

(Van der Meer, 2007). Table 14 summarises the different business objectives of co-development.

Table 14. The Different Business Objectives of Co-Development

Objective	Business requirement	Implication for co-dev design
Increase profitability	Lower cost	Increase volume to spread fixed costs; partner for less critical components
Shorten time to market	Incorporate already-developed components or subsystems	Seek partners with proven capabilities
Enhance innovation capability	Increase the number and variety of front-end technologies	Create strategic research partnerships with universities, research labs
Create greater flexibility in R&D	Share risks with partners	Develop research partnerships in bottleneck areas
Expand market access	Broaden the pathways to market for products and services	Leverage partner's complementary R&D to tailor offerings to new markets

Source: Chesbrough and Schwartz, 2007.

2.3.5. Tools and Technologies

Information and communication technology (ICT) is an essential enabler for the transfer of distributed sources of information in the OI process. There are a variety of environmental, strategic and economic factors that drive the adoption of open approaches to developing and commercialising technology.

Many firms are not exhausting the potential to collaborate with actors outside their organisation to identify means to create and later commercialise their ideas (Chesbrough, 2003b). ICT is important in expanding the competence of firms to collaborate across different geographic and organisational boundaries (Pavitt, 2003). ICT maintains a variety of value-adding services, such as web services, enterprise resource planning and customer relations management. The benefits associated with ICT innovation are based on speed, processor power performance, connectivity and interfaces in combination with cost reduction in equipment, and open computer system architecture (Dodgson et al., 2006). There are a variety of studies on OI discussion technologies, tools, and processes and they are divided into three main types: coordinating/aggregating, liberating and allowing/including.

2.3.6. Ambidexterity in Open Innovation

OI can be beneficial for firms (Chesbrough, 2003a). However, it is still unknown whether OI can improve the firm's ambidexterity. Ambidexterity is the ability to pursue opposite, competing and contrasting objectives such as exploration and exploitation, as well as radical and incremental innovation (Tushman and O'Reilly, 1996). Ambidexterity is seen as one of the main determinants of a firm's long-term survival; hence, the effects of ambidexterity are relevant to OI for research and practice. The interactions between different actors (users, customers, suppliers, universities and innovation intermediaries) can have a positive impact on ambidextrous behaviours of firms (Faems et al., 2005). However, the determinants or conditions under which it is achieved are still unclear (Drechsler and Natter, 2012). In markets that are defined by rapid technological change, short product life cycles, more informed and demanding customers, the firms' success relies on two innovation outcomes: novelty and efficiency (Alegre et al., 2006). On the one hand, efficiency is seen as a reduction in innovation risks, and a decrease in costs and time to market. On the other hand, firms need to consider the introduction of new or improved products, services or processes and its effect on the opening up of new markets. According to Lavie et al. (2010), the two outcomes are critical, because efficiency can produce revenue that is devoted to future innovations (Lazzarotti et al., 2017).

Open innovation refers to the cohabitation of inward and outward open innovation in the innovation operations of businesses (Chesbrough, 2003; Cassiman and Valentini, 2016; Dahlander and Gann, 2010; Enkel et al., 2009; West et al., 2014). Research has indicated that inbound and outbound open innovation complement one another (Cassiman and Valentini, 2016; Enkel et al., 2009). Firms' open innovation attempts are better when they work together. Open innovation is a way to look at both inside and outside opportunities for new ideas to be used (West and Gallagher, 2006). According to Idrissi Fakhreddine and Castonguay (2019) and Laursen and Salter (2006), open innovation gives businesses more opportunities by connecting them to different channels. They also say that tailored innovation skills are needed for innovations to work (Hwang and Lai, 2021).

As new technologies are increasingly commercialised in international markets, attention to the use of external sources of knowledge in firms has increased (Chesbrough 2003b). The literature suggests that firms should use internal and external sources of knowledge to accelerate innovation (Cassiman and Veugelers, 2006; Chesbrough, 2006). It is critical to balance both sources of knowledge (internal and external) to prevent over- or under-investment in R&D, but also to exploit new horizons efficiently (Cohen and Levinthal 1990; Narula 2001). This reflects models proposed in the OI paradigm, which advocates that much of the favourable knowledge for the creation of new products and services exists outside the boundaries of the firm (Gassmann and Enkel 2004). Consequently, exploiting external sources of knowledge accounts for new sources of competitive advantage

(Gassmann et al. 2010; Del Giudice et al., 2013). Firms should collaborate with other actors in markets to improve their innovativeness and sustain their international competitiveness. This “openness” is one of the critical factors to access new ideas, knowledge, skills and technologies from outside. The collaborators within these ecosystems are possibly companies, universities, public and private research centres and citizens, sharing complementary resources, infrastructures, knowledge and technologies (Ferraris and Santoro, 2014; Carayannis et al., 2015).

While both types of innovation are essential, open innovation directly impacts how well businesses use and develop their existing knowledge and skills or how well they learn and develop new skills. Firms that participate in open innovation activities look for and get more resources from outside sources for their R&D and innovation. In contrast, when a company's innovation is limited, it often has to develop its ideas with limited resources or miss out on other opportunities (Koput, 1997;Chesbrough, 2003b; Wang et al., 2015). Open innovation affects more than how firms use and extend their existing resources. It also affects the creation and development of new resources and new ways of competing (Cassiman and Valentini, 2016; Chesbrough, 2003a; Laursen and Salter, 2006; Love et al., 2014; Chesbrough and Appleyard, 2007) (in Hwang and Lai, 2021).

The competence of firms to identify the benefits of new, external knowledge, adapt it, and use it for commercialisation is critical to the firm’s innovation capacities (Cohen and Levinthal, 1990). External knowledge is essential for improved innovativeness and reduces time to market (Enkel, Gassmann, and Chesbrough, 2009). OI consists of inter-organisational life cycles. An inter-organisational view assumes that innovation is the result of the interactions between miscellaneous and bilateral actors that comprise a network in which some actors deal with exploration and others with exploitation. Ambidexterity is not understood at the organisational level, but at the level of a network of organisations (Ferrary, 2011). It means the ability to simultaneously handle contradictory knowledge, exploit current competencies and explore new domains with equal dexterity (Lubatkin et al. 2006). March (1991) associates innovation and knowledge management with exposing tensions around exploitation and exploration. Both practices bring about the combination of knowledge, both existing and levering distributed knowledge (Taylor and Greve, 2006). According to Wadhaw and Kotha (2006), exploitation requires efficiency to mobilise current abilities, and exploration calls for experimentation. For Tushman and O'Reilly (1996), both practices are vital to long-term performance and successful product development (Sheremata, 2000). According to Lubatkin et al. (2006), handling exploitation and exploration are fundamentally rich on conflict (Andriopoulos and Lewis, 2009).

The literature differentiates between radical and incremental innovations. While the former requires products, goods, services, processes and organisational innovation that are new, and is considered the most beneficial due to its contribution to sustainable competitive long-term advantage (Song and Thieme, 2009; Sorescu et al., 2003), the latter makes minor changes to existing products (Jansen et al., 2006) and is critical for short-term competition (Connor, 1999; O'Reilly and Tushman, 2013). Due to rapid technological change and globalisation of markets, firms can no longer innovate internally on their own, so they need to become ambidextrous, using both internal and external sources of knowledge (Chesbrough, 2003b; Chesbrough and Bogers, 2014). Another fact is that explorative and exploitative learning compete for the same scarce resources of the firm (Benner and Tushman, 2015; March, 1991). The principles of OI (Chesbrough, 2003a) have gained traction, to complement internal knowledge creation efforts and to manage the tensions in learning and innovation (Carayannis et al., 2017; Parida et al., 2012). Accordingly, the tensions between radical/incremental and internal innovation and external/internal sources of knowledge have become increasingly relevant for theory and practice (Ardito et al., 2018).

Ambidexterity in the OI process of firms is paradoxical from the point of view of the externalisation of the firms' activities through outsourcing or by establishing alliances (Baden-Fuller and Volberda, 1997, Holmqvist, 2004, Lavie and Rosenkopf, 2006, Rothaermel and Deeds, 2004). However, the literature on organisational ambidexterity has mostly paid attention to the way firms address exploitation and exploration internally. Eisenhardt and Martin (2000) outlined the dangers of discontinuity, when firms source all their knowledge internally. Rosenkopf and Nerkar (2001) presented empirical evidence that exploration beyond organisational boundaries is more beneficial for firms. Puranam and Srikanth (2007) depicted the difficulties faced by firms trying to renew their knowledge through the acquisition of innovative firms. The difficulty of such integration lies in the choice between leveraging existing knowledge or the capacity for ongoing innovation by the target firm. It is understood that new knowledge is partly responsible for the rearrangement of existing knowledge. As Kogut and Zander (1992, p. 384) argue, "combinative capabilities" are the firm's ability "to synthesise and apply current and acquired knowledge." Henderson and Cockburn (1994, p.66) define "architectural competence" as "the ability to access new knowledge from outside the boundaries of the organisation and the ability to integrate knowledge flexibly across boundaries within the organisation." Hence, ambidexterity is compelled by internal and external knowledge processes as well as integration across boundaries (Raisch et al., 2009).

Ambidexterity is also present in the contrasting cultures and beliefs about the degree of openness of the innovation process in firms (Herzog and Leker, 2007). OI describes the dialogue and open nature of innovation and new business development. It reflects the idea that to successfully innovate requires control of all stages, from the generation of ideas, development and production to marketing, distribution, service, and financing. Yet the dimensions of control are shifting. New ideas do not necessarily need to start internally, and the commercialisation does not need to be done by the firms' activities (Chesbrough, 2003a). Possibly firms do not have all the necessary skills to accomplish the entire task from beginning to end. Contrasting closed innovation processes with OI enables the use of external ideas combined with knowledge within internal R&D. Commercialisation uses out- as well as in-house channels to markets that might even differ from the current businesses of the firm, through external pathways such as licensing arrangements or spin-offs (Chesbrough, 2003a). That been said, OI consists of two different pillars: external/internal technology sourcing and internal/external technology commercialisation (Herzog and Leker, 2007), where knowledge flows through a semipermeable corporate membrane, leading to inbound and outbound OI (Chesbrough and Crowther, 2006). According to Herzog and Leker (2007), OI means "systematically encouraging and exploring a wide range of internal and external sources for innovation opportunities, consciously integrating that exploration with firm capabilities and resources, and broadly exploiting those opportunities through multiple channels" (Bröring and Herzog, 2008).

2.3.7. L*ink between OI and Free and Open-Source Software*

ICT provides the digital infrastructure for the inexpensive, rapid and secure storage and transfer of information and data. Similarly, to other information commodities, intellectual property (IP) is crucial in the software industry. However, software firms are increasingly developing free and open-source software, and OI offers a suitable lens through which to examine how firms exploit and explore open-source software. Over the last two decades, firms have been collaborating to develop business applications such as Linux operating systems, Firefox internet browser and the Apache Web Server. To take advantage of external sources of knowledge, firms need firstly to find such innovations, manage the firms' absorptive capacity and couple such spill-over with firm-specific internal innovation to develop products and services suitable to the firm's specific needs. Open-source software consists of source codes that can be edited and shared (Perens, 1999). Open-source software combines different collaborative traditions dating as far back as the 1970s, including BSD Unix and the free software movement by Richard Stallman (McKusick, 1999; Stallman, 1999). Open source refers to particular sets of IP policies and software development methodology, where collaborators are geographically dispersed yet work together on a piece of software (West and O'Mahony, 2005). Initially, programmers were not motivated by financial aspects; however, as the movement grew and the quality of the software improved, paid professionals joined the development and started selling related products and services. Similarly, to other forms of OI practices, ICT is said to be one of the main drivers of free and open-source software (Scacchi, 2004). OI can be described as the systematic stimulation and the exploration of a variety of internal and external sources for innovation opportunities, purposely accommodating the ability to explore a firm's capabilities and resources, and exploiting the opportunities through multiple channels (Cohen and Levinthal, 1990). Considering this definition, the OI paradigm goes beyond the use of external sources of innovation such as customers, rivals and universities (von Hippel, 1988). It differs from the use, management, and employment of IP, as seen in the technical and research-driven generation of IP (West and Gallagher, 2005).

Open-source software has come into the spotlight as the protagonist in the use of knowledge in a network structure (Dedrick and West, 2004; Lerner and Tirole, 2002; O'Mahoney, 2003; von Hippel, 2005). Other studies have discussed the use of alliances (Gerlach, 1992) and the construction of networks by firms as different ways to actively integrate such external knowledge into the innovation processes of firms (Gomes-Casseres, 1996). Powell et al. (1996) analysed the cost and benefits of networks for innovation, and Dyer (1996) used the concept of networks and keiretsu formation in the automotive industry. Recent studies have investigated the adoption of alliances between technology-based industries and companies (Nooteboom, 1999) and the rise of intermediate markets in particular industries (Arora et al., 2001). These intermediate markets influence the enticements for innovation and the way new technologies and firms enter an industry (Gans et al., 2001).

A major development in the past ten years is the increasing number of studies of the role of communities outside the boundaries of firms in creating, shaping and disseminating technological and social innovations. The role of the user in the process of innovation has been known since von Hippel (1988). Still, it was free open-source software that accentuated the vital role of communities for innovation. The expansion of the open-source software movement, and new collaborations, problem-solving and IP practices, have shifted the attention of academia to the phenomenon of "community" and its impact on innovation theory and practice. Parallel to the rise in studies on the role of communities in the innovation process, some researchers concentrate on ''OI'' (Chesbrough, 2003c), looking at topics such as inter-firm cooperation and development of an ecosystem of firms, sharing technologies and trading IP, within a given industry or sector (West et al., 2006). But research on the importance of communities (non-firm actors) is rather rare in the OI literature. What makes open-source software innovation so interesting in the area of OI is that, by its very nature, community-based innovation takes place outside the firm's boundaries. Therefore, communities and their role in the innovation process correspond to and extend the firm-centredness of OI (Chesbrough, 2003a; Chesbrough et al., 2006; Gassmann, 2006). The literature on communities of practice suggests that knowledge transfer between community members is the main precondition to organisational learning, especially for communities involved with the organisation. Several case studies look at user innovation, such as peer-to-peer software development (Lakhani and von Hippel, 2003) and sporting goods (Franke and Shah, 2003). According to von Hippel (2001), assimilation and dialogues with user communities promote innovations by the imitation and the extension of other user innovations (West and Lakhani, 2008).

OI can be understood as a "private-collective" innovation model. Contrasting with the private investment model of innovation based on Schumpeter's ideas, the free revelation of inventions, discoveries and knowledge is also described as OI (von Hippel and von Krogh, 2003/2006). The spill-over can occur in two ways: through compensation (licensing) or without compensation (open-source initiatives). The process of OI, or the opening up of the innovation process, has gained popularity in different sectors of the economy. OI is now a reality in industries such as software, electronics, telecom, pharma and biotech (Chesbrough, 2003c). The impact of open source on the software industry is so strong that even monolithic firms such as SAP, Apple and Microsoft are now increasingly developing shifting resources to build on decentralised research labs to develop their absorptive capacity for outside-in innovation processes. Today, most open-source projects are run by large firms and developed by paid programmes (Dahlander and Wallin, 2006). The majority of the communities consist of a firm's employees contributing to documentation and quality control of codes. OI communities are often users of the output too (von Hippel, 2005); their outputs improve the projects they take part in by achieving the firm's own goals and ideals. Therefore, it is crucial for a firm participating in an open-source initiative to be able to profit from their activities through deployment, hybridisation, complements, or self-service; otherwise, the projects become unsustainable over time. Although many organisations have acquired a great deal of knowledge of these communities, significant progress is often achieved when the major part of the code is written by programmers on the staff (Chesbrough and Appleyard, 2007).

2.4 OUTBOUND OPEN INNOVATION

2.4.1. External Technology Commercialisation

OI has gained considerable attention in the technology management research literature since Chesbrough (2003a), but it was mostly discussed from the viewpoint of external knowledge acquisition in practice and theory (Laursen and Salter, 2006; Lettl et al., 2006; Piller and Walcher, 2006). The external exploitation or outbound OI has not found the same level of interest in implementation and research (Enkel et al., 2009). The most cited and known case studies of outbound technology transfer have mobilised selected industry giants (Arora et al., 2001; Grindley and Teece, 1997; Kline, 2003; Sullivan and Fox, 1996), mainly discussing the out-licensing of technologies (Rivette and Kline, 2000). External technology commercialisation has generally been seen as an ad-hoc activity, and most firms did not consider external technology commercialisation as a systematic business strategy, but rather on a case-by-case basis (Tschirky et al., 2000). Over the last 40 years, technology transactions have expanded substantially. This is because firms are now actively taking advantage of the use of external acquisition of technologies/knowledge as part of their business strategy (Granstrand et al., 1992). The interest in external exploitation of knowledge has increased and has seen continuous growth in line with the growth of knowledge markets (Arora et al., 2001). The commercialisation of technologies includes far more than the simple act of selling their own products, processes and services. It involves other approaches such as licensing, patent selling, technology spin-offs, and technology-induced strategic alliances (Escher, 2005). Chesbrough (2006) explains the external exploitation of knowledge as “the use of purposive outflows of knowledge to expand the markets for external use of innovation”.

ETC refers to the selling or licensing of internally produced technology to clients outside of the company. The transfer can occur between companies in the same sector or companies in other industries (Chesbrough 2011; Easterby-Smith et al. 2008; Enkel et al. 2009; Fosfuri 2006). ETC should not only be considered through licensing but also in the form of collaborations, partnerships, joint ventures and the formation of spin-offs. These alternatives can be classed as strategic or technological alliances (Bouncken et al. 2016; Kim and Choi 2014), licensing (Gleave and Feess 2016), and technology sales (Birkenmeier 2003; Ford 1985; Granstrand 2000). External knowledge exchange and access, which result from interaction with possible technology purchasers and their technological know-how, encourage internal knowledge development, such as through novel technology combinations (Cheng and Huizingh 2014) and facilitate the introduction of new products and services to the market (Enkel et al. 2009; Faems et al. 2010). A too-open approach toward ETC might result in more significant profit setbacks, mainly when technology competition is discarded carelessly (Kline 2003; Laursen and Salter 2006). Too strong a focus on ETC may yield short-term gains in terms of licensing fees but negatively impact the long-term value offered to consumers by losing sight of their requirements (Helm et al., 2019).

Escher (2005) also consider the literature on outbound OI not well explored. However, there have been some attempts at constructing an explicit process description with distinct steps (Escher, 2005; Gassmann and Enkel, 2004). External technology exploitation is not the core business of most industrial firms (Teece, 1998; Davis and Harrison, 2001). Many firms do not fully benefit from their external technology exploitation capabilities (Fosfuri, 2006; Sirmon et al., 2007). Outbound OI considers different options for the strategic externalisation of a firm’s technological assets. While monetary incentives are dominant in most technology transactions (Rivette and Kline, 2000; Davis and Harrison, 2001), non-monetary factors have gained in relevance (Arora and Fosfuri, 2003). There is an established literature on the objectives or incentives for external technology exploitation (Teece, 2000; Escher, 2005; Koruna, 2004). The primary assumption for the diffusion of technology through external commercialisation is based on the premise that a firm needs to rapidly obtain a more extensive infiltration of its technology in its target markets, with the ultimate goal that the technology eventually becomes standard (Conner, 1995; Ehrhardt, 2004; Gassmann and Enkel, 2004) or the firm wants to expand to entirely new markets (Adam et al., 1988; Koruna, 2004).

According to a study carried out by Helm et al. (2017, pp. 342) on the automotive industry in Germany, external technology commercialisation is not always beneficial for firms. For the authors, “the more openness, the better” is not always right due to the costs and risks associated with external technology commercialisation. The authors continue to say that it is very optimistic to believe that external technology commercialisation is always the right approach to firms’ problems when managing their unused knowledge. External technology commercialisation is heavily dependent on the strategic, operational and administrative context of firms and the right trade-off between gains and costs. The authors demonstrated that the extent of maturity of the internally unused technology might play a critical role. Table 15 summarises the pros and cons of external technology commercialisation based on Helm et al. (2017).

Table 15. Pro and Contra External Technology Commercialisation

Advantages and Opportunities	Disadvantages and Threats
Good impact on operational outcomes attributed to higher revenue rates and earnings	Resource intensive The need for specialised human resources Potential conflict between external technology and core business
Greater financial scope of action for the amortisation of R&D costs	Potential loss of focus on core business Lack of control of externalised knowledge
Promotes the revenue quantity of firms and reduced sunk costs	Fostering non-core business innovation of other players Fostering opportunistic behaviour of competitors Incurring more costs by developing external technology commercialisation
Higher pace of innovation	The need to manage increased complexity

Source: Helm et al., 2017.

External technology commercialisation (ETC) has grown in importance in both corporate management practice and academic literature. ETC is frequently acknowledged as highly essential, and it is commonly seen as a vital competency of businesses. Yet, research has mainly concentrated on the potential and benefits of ETC while risks and costs have been overlooked or minimised. ETC may not only result in increased operating expenses, but it may also necessitate more human resources. From a strategic standpoint, ETC may result in the loss of competitive advantage due to the revelation of internal information or the inefficient deployment of research and development (R&D) resources, negatively impacting a firm's performance. The ETC literature has primarily concentrated on pointing out the possibilities, the benefits and especially the monetary effects of ETC, highlighting it as a "must-do" in a firm's action portfolio, while dangers have tended to be overlooked (Frishammar et al. 2015; Hung and Chou 2013). Despite a shift towards more attention to the unfavourable elements of OI (OI) in general by devoting effort, for instance, to the "paradox of openness" (Laursen and Salter 2014), a differentiated solution for companies that want to enhance their competitive position has not yet been provided (Helm et al., 2019).

2.4.2. Desorptive Capacity in Outbound OI

The concept of desorptive capacity combines two strands of literature. It merges the theory of absorptive capacity with dynamic capabilities. Desorptive capacity is understood as a category of dynamic capabilities as a firm intentionally develops, broadens and changes its resource base (Helfat et al., 2007). According to Ziegler et al. (2013), desorptive capacity is defined as the firms' ability to externally commercialize their internal knowledge, in other words firms external knowledge commercialisation capacity.

Information asymmetry acutely influences the technology market and is bi-directional: firms managing the licenses lack the know-how on technologies, and the licensors lack knowledge about the market potential of the technology. Desorptive capacity is a process of learning that is based on exploitation and exploration (March, 1991). One way firms can develop their desorptive capacity in outbound OI is by understanding their technological trajectory (Dosi, 1982), as innovation is path-dependent and a cumulative process of incremental problem-defining and problem-solving activities (Rosenberg, 1982; Dosi, 1982; Garud and Karnoe, 2002). According to Cantwell (2004), some firms are still developing their competencies in technology fields that they had initially started over 100 years ago. Thus, a firm's current technology portfolio is a development of its past problem solving (Hu et al., 2015).

The term "desorptive capacity" has received less attention in academic discussion than its counterpart "absorptive capacity". Both concepts discuss the cognitive barriers/facilitators of knowledge transfer (Le Masson et al. 2012) and are critical of firms' capacity to develop new knowledge. Nonaka and Takeuchi (1995) argue that knowledge is the result of the interactions between transferor and recipient, as it banks on the levels (absorptive and absorptive capacity) of both. If firms need to understand how to decode and use the ideas of others, it is also critical for firms to transfer knowledge. If the two involved firms are not able to align the two processes and achieve the "right combination," science-based collaborations are likely to be very rare (Dell'Anno and del Giudice, 2015).

Outside sources of knowledge are critical for the innovation processes of firms (Cohen and Levinthal, 1990). Most importantly, prior experience is one of the essential factors to perceive and assess the likely use of external knowledge. Organisations may benefit from newly obtained knowledge as long as they can combine it with their existing knowledge (Cohen and Levinthal, 1990; Todorova and Durisin, 2007). Fosfuri (2006) believes that technology alliances or technology licensing are excellent examples of desorptive capacities, because firms can benefit from licensing revenues and the lessons learned related to this type of outward knowledge transfer (Chesbrough (2006) (in Müller-Seitz, 2012). Desorptive capacities are said to be a new concept within the literature on knowledge diffusion, and there is little research available for theory reification or application (Florén and Frishammar, 2012). Firms that want to desorb knowledge from their partners (suppliers, customers and competitors) need to make sure that the dialogue is harmonious with their respective partner's absorptive capacities, and even provide support with the task if necessary (Pagellet al., 2010; Lee et al., 2014). According to Modi and Mabert (2007), desorptive capacities can be a very beneficial process, as knowledge transfer can improve both firms, the one transferring the knowledge and the one receiving it (in Meinlschmidt et al., 2015). Bianchi et al. (2011) also acknowledge the positive impact of the support for technological absorption of partners. Outbound OI plays a major role in characterising desorptive capacities. Yoo (2011) has listed four factors to measure a range of inside-out innovation: knowledge, planning, transfer, and follow-up (Kim et al., 2014).

Based on the capability-based framework for OI and institutional theory, Xie et al. (2020) investigate the mediating effect of desorptive capacity on the relationship between innovation intermediaries and outbound OI in China. The authors conducted a survey data from 217 Chinese manufacturing businesses. The result showed that innovation intermediaries affect outbound OI through desorptive capacity through two parallel channels: identification capacity and transfer capacity. Their work relates outbound OI to innovation intermediates, revealing the critical role of this external element in promoting outward knowledge transfer. The authors described how innovation intermediaries impact outbound OI by adding desorptive capacity, a key but seldom explored capability in OI activities. They focused on the two sub-dimensions of desorptive ability, identification capacity and transfer capacity, and demonstrated their parallel mediation functions. Additionally, their study identifies the boundary condition under which innovation intermediaries influence outward OI through desorptive capacity (Xie et al., 2020).

2.4.3. The Concept of Outbound OI

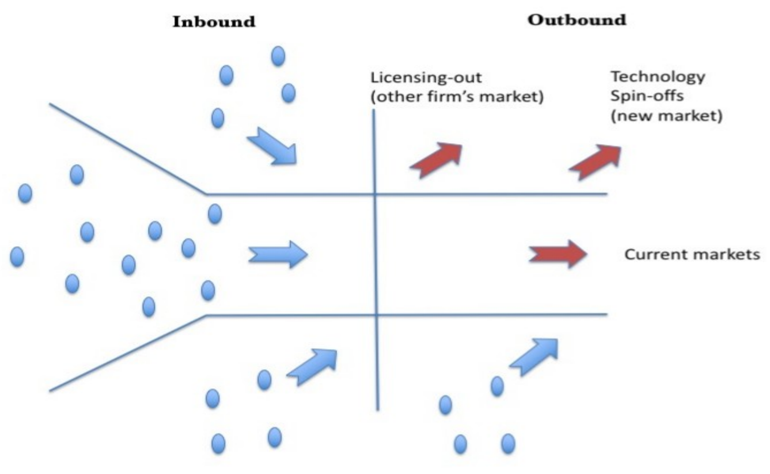
OI adheres to a long-standing tradition in studies on processes of innovation. It has its roots in Schumpeter’s (1934) comparison of the entrepreneurs and the entrenched incumbent firms. The literature on the history of business highlights the different markets for innovation that preceded the corporate R&D departments and the implementation of IP laws (Lamoreaux and Sokoloff, 2001). According to early historical studies on business, R&D activities arose from the necessity in many industries to maintain and improve production activities (Chandler, 1990). As processes and products are unique for each organisation, investment in R&D is particular to each firm. Outbound OI moves away from this rigid organisational thinking and advocates that rather than depending on internal channels to markets, firms should scan for external firms with more suitable business models to commercialise a given technology (Chesbrough and Crowter, 2006).

Outbound OI consists of the flow of internal knowledge toward external environments. Outbound OI is defined by firms leveraging the outcomes of their R&D projects by enabling other firms to take advantage of the firms' own knowledge or the creation of spin-offs (Dahalander and Gann, 2010; Enkel et al., 2009; Huizingh, 2011). Thus, outbound OI deals with: (1) the selling or out-licensing of IP to other firms' markets; (2) creating spin-offs out of internal technologies in new markets; (3) commercialising internally developed technologies in the current market (Chesbrough, 2003/2006; Enkel et al., 2009). OI scholars (Enkel et al., 2009; Huizingh, 2011) claim that by pursuing outbound innovation, organisations may generate alternative revenue streams by exploiting the results of their internal innovation processes (Natalicchio et al., 2017). The three types of outbound OI practices are illustrated in Figure 1.

Outbound OI is contemplated as an additional strategic resource in the toolbox of innovative organisations. Firms can generate value from their innovation without committing many additional resources. In practice, outbound OI can take a variety of legal arrangements, e.g., out-licensing agreements, spin-offs, or technology sale. Outbound OI suggests the use of external pathways to commercialise an innovation (Mortara and Minshall, 2011) and advances both financial and non-economic gains for organisations (Arora et al., 2001; Grindley and Teece, 1997), or “the use of purposive inflows and outflows of knowledge to accelerate internal innovation, and expand the markets for external use of innovation, respectively” (Chesbrough et al., 2006).

Outbound OI encourages firms to be more open in their activities. It is based upon the familiar model of vertically-integrated industrial R&D promoted by Freeman (1974) and others (Chandler, 1990; Mowery, 1990; Pavitt, 1991). Under the paradigm of outbound OI, large organisations do not discard Freeman’s ideas, but rather amplify their conventional R&D practices with composed outflows of internal technologies, pursuing new markets through outbound processes such as licensing (Chesbrough, 2006).

Figure 1. *OI funnel*



Note: Based on Chesbrough (2006).

It is understood in the OI literature that outbound OI is less practised among managers, either because outbound OI is not necessarily associated with cost-reduction measures, or because most firms are users instead of creators of technologies. Additionally, during its first decade, outbound OI was mainly understood as out-licensing patented IP (Chesbrough, 2003) and not necessarily associated with strong appropriability of technologies (Christensen, 2006; West, 2006; West et al., 2006; Dahlander and Gann, 2010). This preconception can be associated with the fact that firms taking advantage of OI were mostly interested in leveraging inbound innovations from non-profit or government labs that intentionally allowed their technologies to spill over without monetary compensation, such as the Linux operating systems (Chesbrough, 2003b; West and Gallagher, 2006; Dahlander and Gann, 2010; Henkel et al., 2014). Research on OI has expanded substantially in the last decades and has gained in both quantity and quality. The original conception of outbound OI extended, shaped and influenced innovation studies and other areas of social science research (West et al., 2014).

On average, outbound OI practices are seen as less critical by firms. However, there has been a slight increase in recent years and, according to a study conducted by Chesbrough and Brunswicker (2014), there has been a positive change since 2008. The authors have reported some improvements in practices such as joint venture activities, the sale of market-ready product ideas to others and participation in public standards development. But outbound OI still has a long way to go when it comes to IP donations and spin-offs. Other authors have investigated the role of “free revealing” (Harhoff et al., 2003). While “free revealing” is well discussed in the free and open-source software literature, case studies show that large firms in other sectors are also trying it out (Tapscott and Williams 2006). In other words, the two pieces of research have confirmed that large corporation are both information takers and givers (Chesbrough and Brunswicker, 2014).

There was also some effort in the outbound OI literature to explore strategic peculiarities of the outbound processes to unrelated diversification within a technology-driven company. Salampasis and Mention (2019) claim that outbound OI activities produce proactive and reactive entrepreneurial initiatives that play a catalytic role in the expansion, growth and value increase of ecosystems. Outbound OI activities underpin a dynamic partnership with ecosystem stakeholders that offers new opportunities for business projects leading to shorter periods of ideation, prototyping, growth and commercialisation, both in linked and unrelated markets, increasing discovery and exploitation capacity. The types of activities will further internal, external and ecosystem-level market resilience, while eventually removing technical, information, organisational and cultural barriers. According to Salampasis and Mention, outbound OI can be seen as the drive and incentive towards knowledge externalisation and technology commercialisation. The effort to externalise monetary and non-financial proposals is a guiding force for firms to open up and look beyond organisational boundaries (Salampasis and Mention, 2019).

Although outbound OI has proved to be beneficial and an efficient strategy for accelerating competitiveness, few studies have focused on it (Inauen and Schenker-Wicki, 2012). Outbound OI is often discussed in terms of technology commercialisation opportunities (Frishammar et al., 2012) or out-licensing (Agrawal, 2005). There is still little evidence linking outbound OI with innovation performance. However, some firms are willing to develop further their outbound innovation capacities through licences and/or through customised solutions for specific target customers or customer groups. Although universities and research institutes are the leading innovation suppliers, various firms are supporting these institutions with critical resources such as specific software or other technological tools and competencies that can help to conceptualise and develop the innovations further.

Additionally, firms may cooperate in particular projects of mutual interest, or even collaborate in joint ventures and consortia (Greco et al., 2015).

When examining the progress in OI in practice, three core elements can be identified, of which one is of relevance here: the process of inside out. The process of inside-out indicates financial (and non-financial) benefits obtained by taking products and services to market, for instance selling IP and multiplying technology by transferring ideas to the outside environment. Companies that have as their core business inside-out processes concentrate on externalising their knowledge and innovations to accelerate ideas to market faster than they would be able to on their own. The main reason for engaging with inside-out processes is the locus of profits or exploitation, through licensing IP and/or multiplying technology, outside the firms' boundaries by transferring products and services to others. When firms immerse themselves into the inside-out processes, they no longer confine themselves to the markets in which they are directly involved. Instead, firms take part in other segments by appropriating licensing fees, joint ventures, spin-offs, etc. Alternate revenue streams generate more overall revenue from innovation (Gassmann and Enkel, 2004). A study conducted by Enkel et al. (2009) reported that only 36% of the firms in their sample use an out-licensing policy to externally commercialise their technologies (Enkel and Gassmann, 2008). The report also shows that only large multinationals are actively engaged with an out-licensing strategy to which they allocate substantial resources. Additionally, there is a gain in awareness of venturing corporate activities (Vanhaverbecke et al., 2008), new business models such as new ventures and spin-offs (Chesbrough, 2007b), and the commercialisation of their own technologies in new markets, called cross-industry innovation (Enkel and Gassmann, 2010).

2.4.4. ***Strategic Applications of Outbound OI***

Since Chesbrough (2003a), firms have increasingly acknowledged that they are not able to hold in-house all the competencies they need, and are compelled to open up their R&D through collaborative processes or the commercialisation of IP rights (Gassmann, 2006; West and Gallagher, 2006). Academic research on OI seems to be overshadowed by case studies or implementation projects within firms (Dodgson et al., 2006), and other survey studies on the ratification and ramification of OI strategies (e.g., Laursen and Salter, 2006). It is understood that OI is beneficial for firms. However, there are no appropriate ways of evaluating the effects of the investments of free versus closed innovation approaches. It is critical to assess OI activities, and there are no such metrics systems available (Enkel and Lenz, 2009). Empirical studies on OI adoption have demonstrated so far that firms are more likely to perform inbound than outbound OI, even if, as a matter of course, every inbound effort should bring about a complementary outbound activity by another firm (Chesbrough and Crowther, 2006). Van der Meer (2007) argues that Dutch firms are hesitant to make use of outbound OI activities (54% outbound versus 74% inbound). Van de Vrande (2009) adopted the approach of technology exploration and exploitation, correlating them with inbound and outbound, and found that less than 30% of his sample engaged in outbound OI, whereas 90% adopted inbound processes. Schroll and Mild (2011) demonstrated that almost twice as much firms from his samples embraced inbound OI over outbound OI. Different studies in specific industries have shown similar results. Chiaroni et al. (2009) and Bianchi et al. (2011) found that more than 60% of OI activities in the biopharmaceutical industry are linked to inbound practices. In the German automotive industry (Ili et al., 2010) only one out of 42 companies were exploiting outbound OI processes. According to Huizingh (2011), inbound OI dominates the empirical studies on OI. The author explains that although many organisations use external knowledge, only a few can provide it. Additionally, there is a strong possibility that either the measurement scales, the respondents, or the samples in these studies are biased (Michelino et al., 2014).

The pharmaceutical industry is highly advanced technologically and consolidates in-depth knowledge of different fields. It is defined by high costs, high risk and long-term perspectives due to heightened regulation. Like other sectors, it has experienced the issue of the decline of R&D productivity (Petrova, 2014). Considering these aspects, the significance of OI has been stressed more than in most sectors. OI is critical for firms to build a range of dynamic skills to leverage their assets effectively, both internally and externally. External exploration has been a subject of study since the 1990s; as organisations changed their emphasis to outbound OI, multiple studies have been performed on desorptive capacity (Hu et al., 2015), which is linked to knowledge exploitation. Most of these studies have focussed on how these dynamic capabilities affect firm performance (Shin et al., 2019; Mazzola et al., 2012) and licensing propensities (Hu et al., 2015; Kani and Motohashi, 2012; Arora and Fosfuri, 2003; Kim and Vonartas, 2006). The study conducted by Lee and Kim (2019) differs from others by not reporting the impact of the dynamic capacities of firms but rather focusing on inter-capabilities analysis and dealing with the determinants of out-licensing inclination (Hu et al., 2015), which are related to knowledge exploration and exploitation (Lee and Kim, 2019).

Outbound OI through out-licensing is a type of inside-out process which consists of opening up the innovation process to external knowledge exploitation (Mortara and Minshall, 2011). Through out-licensing an organisation can gain much more than economic benefits from the commercialisation of their technological knowledge. It obtains strategic non-monetary benefits which include: (1) external knowledge, (2) establishing industry standards, and (3) the opportunity to operate, based on cross-licensing agreements with other firms

(Arora et al., 2001; Grindley and Teece, 1997). The biopharmaceutical industry is hugely relevant for outbound OI (Bianchi et al., 2011a), particularly out-licensing, which plays an essential role in capturing different types of innovation in the pharmaceutical R&D landscape (Allarakhia and Walsh, 2011). Out-licensing links the large gap between internally developed drugs for a clinical trial by large-scale pharmaceutical firms (Paul et al., 2010), and the commercialisation capacities in the sector as a whole. But out-licensing is still challenging for most firms. A recent survey in Europe reported a large gap between the willingness to out-license and the actual conclusion of the deal (Gambardella et al., 2007). Due to the highly complex environment associated with the biopharmaceutical industry, out-licensing and its link to outbound OI deserves detailed research attention (Hu et al., 2015).

Outbound OI can be seen as the exercise of instituting relationships with external firms to which technologies are transferred for commercial exploitation. Based on the concepts of March (1991), outbound OI is associated with "exploitation" of the firm's current knowledge and technologies (He and Wong, 2004) and often happens in the second part of the innovation processes in the biopharmaceutical industry, during clinical tests and post-approval activities. The exploitations take the form of licensing out, spinning out of new ventures, sale of innovation projects, joint ventures for technology commercialisation, the supply of technical and scientific services, corporate venturing investments and non-equity alliances. In the second phase, biopharmaceutical firms are less reluctant to open their boundaries to external actors to exploit the results of their innovation activities to secure faster and far-reaching access to markets. The reason why biopharmaceutical firms only practise outbound OI in the so-called "second or exploitation phase" can be associated with the unique ways biopharmaceutical firms innovate. It is only in the exploitation phase that the clinical tests reach their candidate target groups and the drugs enter a stage of development where the drug can be commercially exploited. Before this, the drug discovery and development innovation processes occur mostly through ‘‘trial-and-error’’ activity, defined by high uncertainty and unpredictable outcomes. In the second phase the development risks drop substantially, and the innovation processes stabilise and become more visible to outsiders. Consequently, opportunities for external commercial exploitation arise and continue. According to Bianchi et al. (2011)’s study, commercial exploitation may be initiated before the end of pre-clinical tests. Correspondingly, other firms can exploit their technologies and outbound their ideas later on in the process (Bianchi et al., 2011).

Outbound OI provides firms with a framework in which any product or innovation process is seen as economically beneficial and can be exploited externally. External commercialisation can be an add-on or even replace internal commercialisation. In this sense, outbound OI indicates external exploitation of internal knowledge, and focuses on the exploitation process across boundaries. As firms consistently perform less outbound OI (Chesbrough and Crowther, 2006; Bianchi et al., 2011; Cheng and Huizingh, 2010; Chiaroni et al., 2009), it can be assumed that most organisations neglect potential benefits (Chesbrough, 2003a; Van de Vrande et al., 2009). Rivette and Kline (2000) have pointed out three main reasons for reluctance of firms: (1) historical reasons; (2) the possibility of using existing relationships; (3) the fear of diffusing relevant knowledge (Rivette and Kline, 2000). A later study by Kline (2003) added another element to this: unwillingness to disclose corporate ‘crown jewel’. But there is some evidence that organisations are willing to change and license out their technologies (Fosfuri, 2006; Granstrand, 2004). Further research will be able to demonstrate whether this trend will persist, and will address other issues such as the consequences of lower costs, shorter time to market and more sales. Cheng and Huizingh (2010) challenge new research to discuss topics such as innovativeness, number of innovations, financial benefits, or non-financial benefits. Rigby and Zook (2002) suggested further studies that would build upon issues such as intermediate benefits, better ways to measure the actual value of innovation or to define the core competencies of a company. This kind of research could lead to more research in outbound OI. Nagaoka and Kwon (2006) led the way and reported strategic benefits of outbound OI, including getting access to new markets and enhancing the firm's technological position (Huizingh, 2011).

In outbound OI, a start-up could act as the provider of technology for large firms and grow into a vital source of innovation (Audretsch, 1995). IP plays a critical role when it comes to protection mechanisms, and decreases the risk of technology misappropriation (Andries and Faems, 2013). This was the case with Isobionics: the start-up was able to outbound its technology to a large corporation and built a lucrative business. A case study of the Airfryer describes how a multinational is capable of profiting from unused technologies by bringing the technology of a start-up to the market. In most cases, start-ups do not have the skills for large-scale production and have no experience with complex distributions channels. Consequently, start-ups are forced to work with large firms to take advantage of the resources of large corporations for successful commercialisation of their technology. These deals can take different forms: (1) licensing out the technology to the larger counterpart; (2) corporate venturing or acquisitions by the multinational; (3) co-development agreements. It is a win-win situation for both firms. While the corporation does not have to invest in R&D for development, the start-up has better access to markets (Usman and Vanhaverbeke, 2017).

Table 16. Strategic use of outbound OI

Implementation Vehicle	Strategic Objective	Mechanism	References
IP Sharing	The technology is strategically developed for the core business activities	It enables all stakeholders to capture value to stimulate inter-firm cooperation	Leten et al (2013)
Free / selective revealing	Evoke contributions from other actors Grow the overall market	It discloses internal knowledge to attract third-party contributors and users	Alexy et al. (2013) Jappensen and Lakhani (2010)
Out-licensing	Access complementary assets Set new industry standards	Uses cross-licensing agreements for knowledge exchange Promotes the large-scale adoption of new technologies	West (2003)
Open collaboration platforms	Accelerate development and commercialisation of complementary innovations	Engages with external actors (users and firms) to spur ecosystem-related innovation	Chesbrough and Garman (2009)

Source: Masucci et al., 2020.

Although recent studies have helped to understand some of the ramification and implication of outbound OI in organisation (Chiaroni et al., 2011; Mortara and Minshall, 2011), there are still many open questions around the underlying mechanisms of outbound OI. Much of the research has focused on the firm-level elements of outbound activities, ignoring the interdependencies with the activity of other firms in the ecosystem and their project-related decisions (Vanhaverbeke et al., 2014). Little progress has been made with analyses of outbound OI at firm level (West et al., 2006; Chesbrough and Bogers, 2014).

Masucci et al. (2020) have explored outbound OI strategies in ecosystems, bringing to light mechanisms used by firms to accelerate technological progress in their complementary activities (Masucci et al., 2020). Table 16 summarises the outbound OI strategies used (based on Masucci et al. 2020)

Smith and Akram (2017), in their study on knowledge externalisation and technology commercialisation in the public sector, identified four interrelated roles that can facilitate public actors' outbound OI abilities. The study was based on case studies in Sweden's public transportation industry. The authors describe how intermediaries may help both innovation seekers and external innovators overcome social and technological hurdles to innovation. Four possible functions of outbound OI in the public sector are outlined based on this capability: broadening the borders of innovation ecosystems, lowering the costs of remote search and data processing, encouraging inter-organisational cooperation and aiding innovators in controlling their innovation trajectory. Their work compares and contrasts these responsibilities with those of intermediaries in private sector OI. In conclusion, the article adds to public sector OI practices and how intermediaries may be utilised to enhance outbound OI. Knowledge of situations when public actors enhance their exploitation capacity by transferring internal knowledge to third parties (outbound OI) or forming long-term innovation partnerships with complementary partners is lacking. Nonetheless, Lee et al. (2012) found that outsider entrepreneurs have been attempting to leverage the value of government data and Smith and Akram (2017) proposed that additional study on outbound OI in the public sector is needed, driven by the new public governance movement's emphasis on public-private partnership. Table 17 summarises the identified interlinked roles.

Table 17. Identified Roles

Expanding the Boundaries of innovation Ecosystems
Decreasing Costs for Distant Search and Data Processing
Fostering Inter-Organizational Collaboration
Assisting Innovation Seekers in Managing the Innovation Trajectory

Source: Smith and Akram, 2017.

There has been also recent efforts by researchers (Mohamad Hashim et al., 2020) on outbound OI policy for exploitation of intellectual creation, design and creativity in Malaysian public universities (MPUs). According to the authors, a substantial amount of intellectual design, invention, and creativity in MPUs remains untapped due to current intellectual property, innovation, and commercialisation regulations. Consequently, the Malaysian research team looked at the goals, applications, and strategies that are thought to be important in the establishment of an outbound OI policy in MPUS for the exploitation of intellectual invention, design, and creativity. The study analysed and proposed appropriate strategies for exploitation of intellectual invention, design, and creativity in Malaysian public institutions based on a comparative analysis with other top international university. The proposals are appropriate for public universities in Malaysia to implement since they were developed after studying the policies of universities in countries with a comparable legal system to Malaysia, namely Common Law. Universities that adopt outbound OI strategies must change their IP, innovation, and commercialisation policies to guarantee that the program is implemented successfully. Similarly, through outbound OI strategies, public research funding organisations are anticipated to alter their research funding policies to allow permissive licensing of unexploited intellectual production, design, and creativity (in Mohamad Hashim et al., 2020)

Another example of outbound OI in the public sector was given by a Chinese researcher team. Zhen et al. (2018) conducted a survey of 2071 R&D contract transactions between Chinese firms and interational multinationals. The study points out internal factors (e.g. organisational capabilities) as well as external factors (e.g. the role of Intellectual Property protection) that influences the transfer of knowledge, which reframing the focus of Policy from the relational dimension of 'Government vs. MNCs' to the transactional dimension of 'Domestic entities vs. MNCs'. Zhen et al. (2018) followed a quantitative approach offering a new prespective of outbound OI within the context of China's Policy. Similalrly, to other countries nearly 40% of multinationals patents are kept unused in China and the impact of outbound OI on performance are positive and negative (Kline 2003; Laursen and Salter 2006; Enkel 2010). Zhen et al. (2018) focused on the positive effects of outbound Oi and highlighted the creation of new opportunities for the implementation of China's Policy in the context of technology transfers to the other countries, e.g. hydroelectric in the Three Gorges Project and locomotive manufacturing technology.

Along these lines, Bogers et al. (2018) have also stressed the role of policymaking for the future of OI as an academic discipline and industrial practice, including outbound OI. The authors pointed to the growth of Initial Coin Offerings (ICO). ICOs generated a lot of public interest in 2017 when cryptocurrency prices skyrocketed. However, these instruments are not well understood, and our regulatory frameworks are far behind in handling the threats that these instruments pose to investors. The continuous growth in innovation ecosystems has significant consequences for antitrust policies, which remain anchored in an earlier model of competition (Bogers et al., 2018).

2.4.5. *Ambidexterity in Outbound OI*

An organisation that is capable of exploiting its knowledge by transferring it beyond its boundaries can produce value through the development or exploitation of additional resources (Teece, 1986). The creation of value banks on the ownership of valuable knowledge (assets), but it is hugely dependent on the capacity of the firm to desorb knowledge to exploit it effectively. One good example of this is the case of IBM. After the initial learning period, the organisation was able to increase its licensing revenue by spotting potential usage and receivers of its technological know-how and making an effort to adequately transfer the knowledge of the firm with support to others. Additionally, IBM was able to expand its knowledge base network. Other firms such as Sulzer Rüti have shown similar results: the firm transferred its technology to a competitor in a joint venture, and Sulzer Rüti was able to acquire new ideas to improve the transferred technology. Although extremely beneficial, organisations and academia alike have paid relatively little attention to this issue (Florén and Frishammar, 2012; Jia and Lamming, 2013; Scherrer-Rathjeet al., 2014; Hu et al., 2015; Meinlschmidtet al., 2016). Desorptive capacity is partner-specific, as it uses the absorptive capacities of other firms (Dyer and Singh, 1998; Wagner, 2012). Schulze et al. (2014) associate desorptive capacity with the coexistence of exploratory and exploitative learning, in other words, organisational ambidexterity. Organisational ambidexterity is the capacity of firms to explore and exploit knowledge (Andriopoulos and Lewis, 2008; Cao et al., 2009; Gibson and Birkinshaw, 2004; He and Wong, 2004; March, 1991; Tushman and O'Reilly, 1996; Yi and Guo, 2014). Although the literature lacks an understanding of how ambidexterity can impact the firms' ability to profit, there have been recent calls for further study (Hu et al., 2015). The idea is that firms with both learning styles can be more successful than others (Tushman and O'Reilly, 1996; Rothaermel and Alexandre, 2009; Tamayo Torres et al., 2014/2017; Wei et al., 2014).

Outbound OI enables firms to govern their technological capabilities better; this, in turn, may explain why some firms are more ambidextrous in terms of their innovation activities. Outbound OI, or outward knowledge flow, results in ambidexterity by firms accessing this to differentiate their knowledge (Hung and Chou, 2008). Firms with high social status enjoy legitimacy, credibility, and respect from others. Through that, these firms grant access to tacit knowledge within specified technological domains (Hu et al., 2005; Chiu, 2014). This type of knowledge assists other firms as a vehicle through which they can improve their existing resources and adapt their innovation processes to accommodate market and technological requirements. The chances of success increase for organisations that commit to exploitative innovation. Additionally, firms do not sell their technical assets or IP to external firms unless their offers are new to markets. Yet, when an organisation licenses out and reveals its innovation, they potentially gain market approval of their inventive internal capacity (Kauppila, 2010). This, in turn, encourages the responsible actors of the internal innovation to collaborate further in R&D activities to improve the depth of their knowledge. A survey conducted by Cheng et al. (2016) in Chinese high-tech firms confirmed that outbound OI enables firms to transfer internal knowledge to external partners, resulting in exploration activities. Most importantly, outbound OI is useful to accommodate the knowledge base of ambidexterity. The point made by the authors is that when firms are capable of coordinating their knowledge between cross-boundary exploration and exploitation, they are also able to manage knowledge rooted in similar internal exploration and exploitation. When identifying the dissimilarities and reciprocal features between internal exploration and external exploitation, organisations will gather valuable experience in determining and handling comparable differences and complementary features in their internal exploration and exploitation activities. The processes of exploitation and exploration enable firms to engage in dialogue with their environments and develop synergistic value from differentiated knowledge linked with exploration and exploitation, facilitating ambidextrous behaviour. For Dai et al. (2019, p. 79) “if a firm learns to exploit the outcomes of exploratory activities externalised to alliances, this can help a firm better integrate its knowledge”.

2.4.6. Neither Invested nor Shared Here

The outbound OI process in firms is a complex activity to put in practice. This complexity can be associated with several elements: (1) the intangible and idiosyncratic nature of disembodied knowledge for which market imperfections prevail (Fosfuri, 2006); (2) the variations of relevant sub-tasks: technical, marketing and legal (Guilhon et al., 2004). It consists of many processes, structures and activities that support the recognition of potential receivers of knowledge, and the management of spontaneous demands from external partners (Chesbrough, 2003a). But, if the “not shared here syndrome” exists, employees are reluctant to exploit knowledge and technologies that cannot be commercialised in their own products and markets. This emotional prejudice influences how employees transfer relevant knowledge, take part in collaboration and mediate contracts with partners, resulting in not sharing knowledge even if it is useless for internal purposes. De Araújo Burcharth et al. (2014) have identified three ramifications for the employee who is not in favour of exploiting ideas outside the organisational boundaries: (1) they ignore potential opportunities for out-licensing of technologies; (2) they need an “excuse” every time a new opportunity arises for external knowledge sharing; (3) they make adverse inferences and judgements about it. These three elements combine to generate a situation where inter-firm knowledge sharing is perceived as harmful, even if this would not necessarily be true, because the “not shared here syndrome” may affect the employee’s behavioural decisions too. To put this in other words, individuals who are reluctant to engage or execute outbound-related activities may even seek to limit themselves, motivated by the idea of not selling out relevant in-house capabilities, resulting in a negative attitude towards exploiting knowledge and technologies outside the normal channels of the company (out-licensing agreements, the setting up of ventures or free/ revealing), and blocking the adoption of outbound OI (de Araújo Burcharth et al., 2014).

2.4.7. Research Gaps in Outbound OI

Although OI has garnered a great deal of interest in the management research literature since Chesbrough (2003a), it has been mostly discussed from the knowledge acquisition (inbound/outside-in) perspective by academia and practitioners. Consequently, the external exploitation of knowledge (outbound/inside-out) has not achieved an equivalent level of adoption in practice and theory (Enkel et al., 2009). It is unfair to say that there have been no efforts at all. But, despite the many achievements in the outbound OI literature, they were limited to selected industry giants (Arora et al., 2001; Grindley and Teece, 1997; Kline, 2003; Sullivan and Fox, 1996), and out-licensing of technologies (Rivette and Kline, 2000).

More recently West and Bogers (2017) argued that the concept of outbound OI has been neglected in theory and practice and there are many opportunities for further research on outbound OI in its role in service and network forms of collaboration such as ecosystems and platforms. Additionally, dynamic capabilities research offers possibilities to explain how businesses gain a competitive advantage through outbound OI tactics that cross organisational boundarries (Christensen, 2006; Teece, 2007; Vanhaverbeke & Cloodt, 2014). Notably, the type of research is characterised by fewer constraints on the external use, modification, and dissemination of project output (von Hippel & von Krogh, 2003). For instance, unlike firm-created proprietary software, free and open-source software projects allow users to alter and distribute their product output with variable degrees of restriction. Thus, users with a high degree of outbound openness may change the product to tailor it to their requirements or redistribute it for other purposes. In this regard, this study focuses on the capabilities of firms to explore and exploit their outbound OI processes in an open-source software big data ecosystem.

As inbound OI is more widely accepted and used in businesses, there are far more inbound OI studies than outbound OI ones. According to Stanko et al. (2017), the number of research publications on outbound OI is less than half that of inbound OI. It supports West and Bogers' (2014) finding of just 50 research articles on the former versus 118 on the latter. Outbound OI appears to be still inadequately represented in research and practice. There are at least two explanations for this research gap in outbound OI. First, organisations are less eager to share knowledge with outside parties than to assimilate it freely. Second, whereas inbound OI may be examined in the early stages of obtaining outside knowledge, outbound OI is typically noticed after completing the lengthier knowledge absorption and integration process. The scarcity of outbound OI research, however, does not imply that the subject is irrelevant. On the contrary, there are many opportunities for further analysis (Cheah and Ho, 2021).

For instance, Verreynne et al. (2020) claims that outbound OI is not always motivated by monetary benefit. Organisations may disclose to improve other elements of their innovation processes (Henkel et al., 2014). As a result, while free sharing occurs less frequently, outbound OI researchers must grasp the complete picture of knowledge transfers between firms. The authors invite academics to investigate the mechanisms that support disclosing from the perspective of an organisation. The current work responds to the call and offers an excellent opportunity to understand outbound OI in a digital service platform, from how firms use their capabilities to outbound their knowledge and commercialise their technology.

Gassmann and Enkel (2004) and Escher (2005) have suggested that outbound OI is not well explored. More recent studies have also called for more research in this area (Mortara and Minshall, 2011; Ziegler et al., 2013; Chesbrough and Bogers, 2014; Smith and Akram, 2017; Bogers et al., 2018, Lee and Kim, 2019; Salampasis and Mention, 2019). Helm et al. (2017) recommended the identification and analysis of additional case studies. For Vanhaverbeke et al., (2014) there is little evidence of outbound OI in service digital ecosystems and little progress has been made (Chesbrough and Bogers, 2014). According to Hu et al. (2015, pp. 47), "outbound OI […] remains a challenge for most firms", and it is not very clear yet why firms cannot improve their outbound OI capacities. For Smith and Akram (2017, pp. 1), "there is a little knowledge about the cases where public actors increase their exploitation capacities by transferring internal innovations to external parties (known as outbound OI)." The lack of commitment can be associated with the simple fact that technology exploitation is not the core business of most firms (Teece, 1998; Davis and Harrison, 2001).

OI has significantly impacted innovation studies and other areas of social science research (West et al., 2014). Still, outbound OI is viewed as less critical for firms. Although, in recent years some progress has been reported, there are still some shortcomings and little evidence connecting outbound OI and innovation performance (Greco et al., 2015). Although recent studies have helped to understand some of the ramifications and implications of outbound OI in organisations (Chiaroni et al., 2011; Mortara and Minshall, 2011), there are still many open questions around the underlying mechanisms of outbound OI. Much of the research has focused on the firm-level elements of outbound activities, ignoring the interdependencies with the activity of other firms in the ecosystem and their project-related decisions (Vanhaverbeke et al., 2014).

Outbound OI proposes the application of external channels to commercialise or advance an innovation (Mortara and Minshall, 2011) and offers both financial and non-economic gains for organisations (Arora et al., 2001; Grindley and Teece, 1997). Nevertheless, it is appreciated in the OI literature that managers engage less in outbound OI than in its counterpart, inbound OI. This is either because managers do not see outbound open as a means for cost-reduction, or it is because most organisations are consumers rather than developers of technologies. Or, because outbound OI has been connected to the out-licensing patented IP (Chesbrough, 2003) and not necessarily associated with strong appropriability of technologies (Christensen, 2006; West, 2006; West et al., 2006; Dahlander and Gann, 2010). Therefore, it is critical to: (1) explore outbound OI processes in firms rather than inbound OI practices; (2) explore other outbound OI processes than out-licensing patented IP, and analyse outbound OI processes in the current market of firms; (3) explore outbound OI processes in technology consumers and creators.

Different industries have contributed with a variety of empirical studies (He and Wong, 2004; Chesbrough and Crowther, 2006; Van der Meer, 2007; Gambardella et al., 2007; Chiaroni et al., 2009; Ili et al., 2010; Paul et al., 2010; Van de Vrande, 2009; Van de Vrande, 2009; Schroll and Mild, 2011; Allarakhia and Walsh, 2011; Michelino et al., 2014) to the outbound OI literature. The studies have proved to be very beneficial to help understand some ramifications and implications of outbound OI in organisations (Chiaroni et al., 2011; Mortara and Minshall, 2011). There has even been a recent effort to understand outbound OI strategies in ecosystems in the energy industries including oil and gas, contributing to the outbound OI literature with examples of mechanisms used by firms to accelerate technological progress in their complementary activities (Masucci et al., 2020). But not many of these efforts have focused on outbound OI processes in free and open-source software service digital platforms ecosystems.

The outbound OI process has received little attention, and the identification of its phases has largely gone unnoticed. The literature has delved into specific ICT technologies that assist with various phases or even individual activities of the OI process. The scientific endeavour has provided a fragmented picture of the kind of assistance that ICT technologies may provide. As a result, the current scientific discussion does not give a comprehensive view that evaluates the linkages between ICT tools as the outbound OI process evolves (Aloini et al., 2019). The literature lacks a framework that analyses the whole outbound OI process, on one hand, and provides a comprehensive perspective of the support that ICT tools may provide to the OI process while supporting the corresponding OI knowledge flows on the other (Dodgson et al., 2006). This is seen as a significant deficiency in the scientific community, and it should fuel the future research agenda on outbound OI (Bogers et al., 2017). Nonetheless, data from the field indicates that companies are becoming increasingly interested in leveraging ICT to assist outbound OI operations. The purpose of this research is to investigate the role and capabilities of Apache Hadoop in supporting the outbound OI processes within firms in their current market (Aloini et al., 2019).

Ambidexterity is the ability to pursue opposite, competing and contrasting objectives such as exploration and exploitation, as well as radical and incremental innovation (Tushman and O'Reilly, 1996). OI advocates that external knowledge is beneficial for firms if they can align the newly acquired knowledge with the firms' business models (Chesbrough, 2003a). Firms that can exploit their knowledge by transferring it beyond its boundaries can produce value through the development or exploitation of additional resources (Teece, 1986). The creation of value banks on the ownership of valuable knowledge (assets), but it is hugely dependent on the capacity of the firm to desorb knowledge to exploit it effectively. Ambidexterity is understood as critical for a firm's long-term survival, and the effects of ambidexterity are pertinent to OI for research and practice. Dialogue with actors (users, customers, suppliers, universities and innovation intermediaries) outside the firms' boundaries are very beneficial and can positively influence ambidextrous behaviours of firms (Faems et al., 2005). The essence and environments under which it may take place are still unclear (Drechsler and Natter, 2012) and have been widely neglected (Florén and Frishammar, 2012; Jia and Lamming, 2013; Scherrer-Rathje et al., 2014; Hu et al., 2015; Meinlschmidt et al., 2016).

Dynamic capabilities are dynamic organisational processes (Teece, 2007). They link the ambidextrous strategies of firms and provide opportunities for outbound OI practices. The work contents that organisational ambidexterity is strategic directives (exploitation and exploration), allowing companies to adapt to and influence their competitive environment. The dynamic capabilities are developed following the ambidextrous strategy of firms, which is displayed by innovating discontinuously, architecturaly and continuously (Lieshout, 2021).

Literature Gap 1:

Outbound OI is usually considered from the perspective of the technology creator, and mostly from an IP-licensing angle. The thesis contributes to the OI literature by adding empirical evidence of outbound OI processes of technology consumers and creators in the firms' current markets.

Literature Gap 2:

There are very few empirical studies to date on outbound OI processes in free and open-source software service digital platforms ecosystems. The work aims at adding empirical evidence to studies on outbound OI in service digital platforms ecosystems.

Literature Gap 3:

Although the literature on ambidexterity is extensive, there are also still some shortages in studies on ambidexterity in outbound OI. Therefore, the study answers to the recent call for more research on ambidexterity in outbound OI processes in firms.

2.5. FREE AND OPEN SOURCE SOFTWARE

2.5.1. The Rise of a Sector

Free and open-source software (F/OSS) has grown from an ideology to a major value proposition for organisations. A substantial volume of software released within the legal framework of open-source documentation are developed and used by many commercial firms. As digital networks enabled by information and communication technologies (ICT) have become increasingly more reliable, faster and omnipresent in most societies, peer production has expanded its remit (Benkler, 2006; Baldwin and von Hippel, 2011). In line with this, hobbyists and practitioners have combined strengths to generate cultural content such as free and open-source software. According to Benkler (2002), peer production is the “quintessential instance” of “commons-based” production and has overturned conventional approaches in software development. Free and open-source software has become a major economic, social, and cultural paradox, stimulating debates for academics and professionals alike (Lee et al., 2009; von Hippel and von Krogh, 2003; Koch, 2004). Benkler (2006, pp. 59) has pointed out that the paradigm change is marked by innovative means of value creation in post-modern societies, “[...] one that should not be there, at least according to our most widely held beliefs about economic behaviour.” Free and open-source software 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 ‘bazaar’ model as separating the two very antagonistic means of software development. On the one hand, the former recognises the innovative nature of releasing ideas early and often, entrusting responsibility to the community, and disclosing all details of the software package. On the other hand, the latter illustrates the rigid and more conventional copyrighted models of commercial software vendors.

Free and open-source software offers a multitude of opportunities to incorporate creative peer networks. As stated by the Free Software Foundation (FSF), free software is associated with four essential freedoms summarised in Table 18.

Table 18. The 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

Source: Sen et al., 2011.

The origin of free and open-source software philosophy can be traced back to the very early days of computing. Back in the 1950s and 1960s, hardware and software were sold together (Weber, 2004; Hars and Ou, 2001), lines of code were shared between software developers and it was common practice to distribute knowledge between experts (von Hippel, 2005; Weber, 2004). Commercial software packages were rare or practically non-existent. If one needed a specific program and/or feature, one had to develop it on their own or pay someone else to do so. It was the prerogative of the research culture of these professionals to freely give and take software codes. This collective software development philosophy is a fundamental component of “hacker culture” (von Hippel, 2005). It is important to keep in mind that computers were only available for some major private and government agencies—they were very expensive, they had to be placed in massive warehouses, were not powerful at all and really hard to take care of. When IBM offered the IBM 701 (the first commercial electronic computer) it cost around $15.000.00/month. The IBM 705 (a commercial version of the IBM 701) was worth circa $1.6 millions (Weber, 2004).

In the 1980s, as software became commoditised and available separately to hardware, Richard Stallman established the FSF (Hars and Ou, 2001) after a simple paper jam on a laser printer offered by Xerox to the Massachusetts Institute of Technology (MIT). Stallman and his colleagues wanted to deal with paper jams by testing and modifying the package in order to improve the codes. Xerox refused them access to the codes of the printer software (Weber, 2004; von Hippel, 2005). Stallman realised that the only way to fight the commodification of the software industry was to use the legalities of copyright to create a system of shared ownership. The FSF built the legal platform for many projects to come, in particular, Linux. The key concepts behind free and open-source software were established in the seminal work, The Cathedral and the Bazaar (Raymond, 1999). Table 19 summarise three possible free and open-source software licenses based on Carver (2005).

Table 19. Free and Open-Source Licenses

Type	Explained
General Public License	The most protective license and guaranties all four freedoms.
Lesser General Public License	It was written to permit a better link to proprietary software. It allows firms to patent some features of the software package. It can be seen as something in-between GPL and traditional proprietary forms of software from the box.
Mozilla Public License	By MPL all source codes need to be openly available. Still, it is possible to combine MPL codes with proprietary codes, as long as the MPL codes are separate from proprietary codes.

Source: Carver, 2005.

Users of F/OSS take advantage of: (1) low cost, (2) development flexibility, (3) low entry barriers and (4) other social and economic benefits. As previously mentioned, users can take advantage of many copyright protections: (1) General Public License (GNU GPL), (2) Lesser General Public License (LGPL GNU) and (3) Mozilla Public License (MPL). Wang (2012) released a well-developed and detailed table of free and open-source software and distinguished the various licensing types as: (1) strong copyleft; (2) weak copyleft and (3). Table 17 summarised the licenses based on Carver (2005) and Table 20 differentiate the various licenses types based on Sen et al. (2011).

Free and open-source software represents innovative services and product shaped by users for users, thereby challenging the major software manufacturers and their research and development strategies. The users themselves develop what they need, rather than relying upon off-the-shelf products that are the manufacturers' response to deceptive market requirements. The modular characteristics of free and open-source software enables users to build upon the work of others through collaborative information and communications technology (ICT) usage.

Table 20. Free and Open-Source Licensing Types

Licensing Types	
Strong copyleft	Most-Restrictive
Weak copyleft	Moderately Restrictive
Non-copyleft	Least-Restrictive

Source: Sen et al., 2011.

Modularity plays a critical role in the process and describes how the components of a given structure can be divided and relinked. It indicates the compactness within constituent elements and the level of the architectural rules enabling or prohibiting the mingling of components within an infrastructure (Schilling, 2000). The foremost innovation around modularity was the creation of coding methods and builders, which (l) enable one module to be coded with little information about the other modules; and (2) permit modules to be congregated and substituted without reorganising the entire architecture. This resource is exceptionally appreciated for the development of huge pieces of code (Parnas, 1972). The practicability of such a modular system is that developers coding in loosely structured groups can create based on the 'private-collective' innovation model. Contributors profit while adding to the common good by 'freely revealing' the source code (von Hippel and von Krogh, 2006). Free and open-source software has grown to be a major player which embodies the forefront of innovation in software development. Software bugs and other limitations are more likely to be repaired in a shorter period of time as disparate users engage with a problem at the same time, 24 hours a day, giving economies of scale in terms of brain power. Different users have different skills and F/OSS offers a platform where these specialists can merge resources. This collaborative and innovative environment pushes forward creativity at a pace that conventional organisations struggle to match. It brings to mind Eric Raymond's (1999) notion of the power of "eyeballs": "Given enough eyeballs, all bugs are shallow."

For software developers with limited skills and/or resources, it is much easier to accelerate service and product development by releasing their codes under F/OSS licence agreements. In some cases, programmers can also learn new capabilities by looking at the source codes produced by others. This concept has a massive impact on society: (1) it promotes cultural exchange, (2) it advances innovative partnership and (3) it can organise clusters of individuals in the margins of societies. However, F/OSS is definitely not a "silver bullet" (Brooks, 1985) and it has its limitations. Table 21 illustrates free and open-source software (F/OSS) timeline to 2000 and Table 22 extends it to 2020.

Table 21. Free and Open-Source Software (F/OSS) timeline

1950-1960	Software source code is distributed without restriction in IBM and DEC user groups
1969	Ken Thompson writes the first version of Unix. Its source code is distributed freely through the 70s.
1978	Donald Knuth (Stanford) publishes TEX as free software
1979	Following AT&T announcement to commercialise Unix, UC Berkeley begins with the creation of its own version of Unix, BSD (Berkeley Software Distribution) Eric Allmann, a student at UC Berkeley develops a programme that routes messages between computers over ARPANET. It later evolves into sendmail.
1983	Stallman publishes GNU Manifesto calling for free software, and establishes the Free Software Foundation
1986	Larry Wall creates Perl (Practical Extraction and Report Language), a versatile programme language used for writing CGI (Common Gateway Interface) scripts.
1987	Developer Andrew Tanenbaum releases Minix, a version of Unix for PC, Mac, Amiga and Atari ST. It comes with complete source code.
1991	Linus Torvalds publishes version 0.02 of a new Unix variant that he calls Linux in a Minix newsgroup.

1993	FreeBSD 1.0 is released. Based on BSD Unix, FreeBSD includes networking virtual memory, task switching, and large filenames. Ian Murdock creates a new Linux distribution called Debian Linux
1994	Marc Ewing forms Red Hat Linux. It quickly becomes the leading Linux distributor Bryan Sparks founds Caldera with backing by former Novell CEO RAY Noorda.
1995	The Apache Group builds a new Web server, Apache, based on the National Centre for Supercomputing Applications' (NCSA's) HTTPd 1.3 and series of patch files. It becomes the dominant HTTP server today.
1998	Netscape not only gives away Communicator 5.0 (Mozilla) but also releases its source code. Major software vendors, including Computers Associates, Corel, IBM, Informix, Interbase, Oracle, and Sybase announce plans to port their products to Linux. Sun announces plans to release the source code for Unix Operating System Solaris
1999	Number of Linux users estimated at 7.5 Million
2000	More software companies such as Novell and real release versions of their products, which run on Linux.

Source: Hars and Ou, 2001.

It goes up against proprietary software products and services in particular critical facets: (1) many software bugs are never fixed due to the lack of guidance and oversight, and certain services and product are not advanced, as there is no economic compensation or incentive; (2) there are no performance guarantees with F/OSS—all product performance liabilities rest with the user; (3) some services and product are not developed to the highest standard—they are just good enough to encourage people to use or download them but they need the support of specialists in order to run adequately; (4) vendor lock-in, once a user has decided to take advantage of a particular infrastructure, can be very difficult and expensive to break.On an organisational level, while proprietary software firms are accountable for their research and development, sales and the extraction of value from their products and services, F/OSS depends on collaboration and ICT. Additionally, F/OSS's source codes, the so-called "crown jewels", are freely available. F/OSS is marked by a paradigm shift in terms of diversity of structure and hiring agreements. This creates a mutualistic relationship between corporations and communities, as well as the creative and commercial angles (Krishnamurthy, 2006). These interactions have most recently been viewed as beneficial resources for organisations. This is another paradigm shift, as these human assets were previously understood as relevant only on an individual scale. Organisations that can effectively apply this model are very likely to reduce the intrinsic hierarchical chain of command and take full advantage of a constantly changing digital economy (Lesser and Storck, 2001). Table 19 illustrates the history of Free and Open-Source Software.

One way for organisations to pursue that is by considering large numbers of indirect connections via ICT, as these offer potential means for stakeholders to overcome the challenges of sustaining direct relationships (Burt, 1992). Inter-organisational liaisons offer two distinct advantages for networks. Firstly, these connections contribute to asset distribution within networks and enable the aggregation of knowledge and organisational equity between firms. Secondly, collaboration within networks also facilitates knowledge distribution, where information about new technological advancements is communicated. These communication channels not only help solve common problems, but also help close windows of failed approaches (Ahuja, 2000). These kinds of relationships are extremely amenable (Piore and Sabel, 1984) and are ingrained in open-ended alliances within social actors—acquiescence is based on reputation rather than standardisation. These exchanges take place across industries, involve big and small participants, and consist of diverse types of relationships (Stark, 2002).

Table 22. Free and Open-Source Software (F/OSS) timeline after 2000

2003	WordPress Release
2004	Ubuntu Release
2006	Apache Hadoop Release
2010	Android Release
2014	Vue.Js Release
2019	Debian Release

Building upon Milgram's (1967) "small world network", and by exploring the connections between creatives in Broadway musicals from 1945 to 1989, Uzzi and Spiro (2005) explain that the more these connections resemble the peculiarities of the "small world", the better the connections are between actors within these networks. Creative outputs are passed through these actors who are connected to other actors outside these networks, extending the radius and reach of collaborative practices. These settings provide a channel for creativity to be spread to other clusters, thereby giving integrity to new contexts developed through this process. There is clear value in linking singular actors through short corridors within comparatively thin networks. Such configurations are also favourable to innovation due to enhanced streams of information and the likelihood of fresh interactions between ability and design. Networks help us to grasp the flow of knowledge and concepts, and the reliant associations between technologies and the circulation of innovation (Steen et al., 2010). Such a configuration pushes innovation forward, as it establishes information 'shortcuts' in large networks (Cowan and Jonard, 2004).

2.5.2. The Economics of Free and Open-Source Software

At the heart of these social-economic exchanges are linkages and communication. Many researchers have devoted a lot of effort towards understanding these principles. They have also contributed numerous theories offering analytical instruments to formulate the different kinds of networks and how these networks transfer tacit and explicit knowledge. Numerous papers have shed light on the central role played by networks in F/OSS (Grewal et al., 2006; Hahn et al., 2008; Singh et al., 2011; Wang and Wang, 2012). On the one hand, these theories suggest different levels of relationships, such as weak and strong ties. These interactions between two or more parties postulate a correlation between the micro and the macro. It is among these associations that we can understand the bigger picture and make sense of small groups (Granovetter, 1973).

As societies are increasingly interacting through ICT, the fashion in which we preserve and approach our social networks is also adapting. The weak ties with our associates are paradoxically much more imperative than the inwardly focused dialogues with our closer keens. In the zone of comfort within our strong ties, we are very likely to use simple 'restricted codes', where much is implied and taken for granted. In interactions with our weak ties, we depend on clearer 'elaborated codes' for the messages to be understood. When developing these ideas, we are exposed to a more creative environment and this process of thinking encourages greater innovation. The weaker the ties we possess, the more we are linked to the world around us. Consequently, we are more likely to obtain critical intelligence about ideas, hazards and openings. Cultures and social systems that have more weak ties are dynamic and exposed to innovation. Systems largely consisting of strong ties are divided and ungainly (Granovetter, 1973).

On the other hand, different levels of ICT usage are suggested in communities and networks of practice. The “network of practice” concept advocates that individuals are driven by shared informal know-how and a desire to accomplish a common goal (Wenger, 1998; Brown and Duguid, 1991). How can this be understood? One way to grasp this issue is by considering that, in networks of practice, most interactions are not face-to-face, but rather enabled by ICT (Lave, 1998; Lave and Wenger, 1991; Wenger, 1998). Such networks join distinctive communities and go beyond physical scope (Brown and Duguid, 2001; Pan and Leidner, 2003; Vaast, 2004, Wenger et al., 2002).

The activities of networks of practice are always shifting because new players replace old ones and the changes demand a revision of interactions within the network. These communities and their agents keep maturing in order to link old traditions with new requirements (Brown and Duguid, 1991). Networks fit particularly well with knowledge-intensive industries where issues are tackled together, as these networks stimulate learning and problem solving (Powell and Brantley, 1992). Network ties are very similar to established partnerships like joint ventures, alliances and R&D partnerships. They can also be compared to customer–supplier relationships, co-market partnerships or the development of complementary products. They reverberate casual partnerships between two or more stakeholders built upon past practice (Simard and West, 2005).

2.5.3. Free and Open-Source Software 2.0

The terminology ‘open source’ was put forward in 1998 to replace the term ‘free software’. Free software was not well accepted by the corporative community due to its business unfriendly connotation. As opposed to free software the open-source movement became very successful, and the up-and-coming OSS 2.0 is particularly commercially driven (Fitzgerald, 2006). This metamorphosis has not happened from one moment to the other but it has rather to be understood as a change over time. Fitzgerald (2006) has developed a comprehensive table with the major differences of F/OSS and OSS 2.0. A copy of this table (Table 23) is presented in the next page of this document and is of indisputable value for the understanding of this phenomena. Table 16 summarises the differences between free and open source and OSS 2.0.

Table 23. Differences Between Free and Open Source and OSS 2.0

Domains	Free and Open Source	OSS 2.0
Development Life-cycle	*Planning* – "an itch worth scratching" *Analysis* – part of conventional agreed-upon knowledge in software development *Design* – firmly based on principles of modularity to accomplish separation of concerns *Implementation* – Code, Review, Pre-commit test, Development release, Parallel Debugging *Production Release* – (often the planning, analysis, and design phases are done by one person/core group who serve as 'a tail-light to follow' in the bazaar)	*Planning* – purposive strategies by major players trying to gain competitive advantage *Analysis & Design* – more complex in spread to vertical domains where business requirements not universally understood *Implementation* – sub-phases as with FOSS, but the overall development process becomes less bazaar-like. Increasingly, developers being paid to work on open source
Product Domains	Horizontal infrastructure (operating systems, utilities, compilers, DBMS, web and print servers)	More visible IS applications in vertical domains
Primary Business Strategies	Value-added service-enabling	Value-added service enabling:

	Loss-leader/market-creating	Bootstrapping Market-creating: Loss-leader Dual product/licensing Cost reduction Accessorising Leveraging community development Leveraging the open-source brand
Product Support	Fairly haphazard – much reliance on email lists/bulletin boards, or on support provided by specialized software firm	Customers willing to pay for a professional 'whole-product' approach
Licensing	GPL, LGPL, Artistic License, BSD, and emergence of commercially-oriented MPL 'Viral' term	Plethora of licenses (85 to date validated by OSI or FSF) 'Reciprocal' term used in relation to licenses

2.5.4. Developers' Motivation

Free and open-source software resulted from the developers' freedom to craft new goods and service (Raymond, 1999; Ghosh, 1998). In this respect, it is important to understand the motivations for these digital craftsmen. Developers' motivation has been of considerable interest in academia and discussed from different angles: intrinsic (Lakhani and von Hippel, 2003; Lakhani and Wolf, 2005), extrinsic or "originating from external rewards" (Lerner and Tirole, 2002/2005), ideology (Stallman, 2002; Stewart and Gosain, 2006), identification with the open-source movement (Hertel et al., 2003; Janssen and Huang, 2008), or a combination of the above (Franck and Jungwirth, 1999; Krishnamurthy, 2006; Roberts et al., 2006), sensitive construct theory in Internet Protocol Television (IPTV) Gomes de Souza (2014) and other technologies.

The different characteristics of the role of the stakeholders in free and open-source projects bring about different social structure. Stakeholders in free and open-source projects develop a community around the project, joined together by their interest in using and developing the software. Members of the projects themselves assign the roles in open-source projects according to their skills and interests rather than being appointed by someone else (Nakakoji et al., 2002). Table 24 summarises the roles in a free and open-source project based on Ye and Kishida (2003).

Table 24. The different Users' Roles in Free and Open-Source Software

Project leader	It is usually the person that has started the project. The project leader dictates the vision and direction of the project.
Core members	Core members guide, manage and overview the development of the project. They are usually members of the projects for relatively long. They have made significant contributions to the development and evolution of the software. Once an open-source project evolves to a certain size, the project leaders may dissolve, and the core members establish a council to take the project to the next level.
Active developers	Active developers are those members that frequently add new features and fix bugs. They are one of the most critical drivers of free and open-source projects.
Bug fixers	Bug fixers fix the bugs they discover, or which are reported by other members of free and open-source projects. They understand in detail the source of the section of the software they are repairing.
Bug reporters	As the name suggests, bug reporters discover and report bugs. Most bug reporters cannot read or programme source codes. They are the testers of the traditional software development model.
Reader	Readers actively use the software and try to understand its underlying principles. Most of the readers are learning how to programme or read the source code to use as a reference to develop a similar piece of software.
Passive users	They are just users of the software. Some users are technically and politically attracted by the philosophy of free and open-source software.

Source: Ye and Kishida, 2003.

The different roles described above can be seen in Linux. Linux is an operating system that was developed as open-source software based on the structure of the UNIX operating system. The Linux itself is made up of various other programmes and modules that are arranged around a kernel. The kernel is responsible for facilitating the access of applications to other hardware resources such as hard disk storage, random access memory, network bandwidth, etc. The kernel is critical for any operating system, including Linux, and contains over two million lines of source code (Bovet and Cesati, 2001). Over the last two decades, Linux is developed by several contributors worldwide (Moon and Sproull, 2002). Many commercial organisations are using and developing Linux (Hertel et al., 2003).

F/OSS has moved away from the initial model of dispersed and decentralised governance and control (Santos et al., 2011). The established peer-based configuration is now being replaced by sponsored, industry-led or industry-involved F/OSS development: (1) sponsored F/OSS is based upon financial injections and/or other kinds of investments from third parties (Capra, 2008); (2) industry-led F/OSS is characterised by a commercial stakeholder calling the major shots (Hou, 2007; Mens et al., 2008; Merlo et al., 2004; Wermelinger and Yu, 2008); and (3) industry-involved F/OSS means that projects are pushed forward by communities but usually have some stakeholders from private or governmental agencies supporting the projects (Capiluppi et al., 2007), resulting in a much-blurred relationship between communities and organisations. F/OSS has also evolved from a horizontal disposition to a more vertical domain of IS(IS) applications (Fitzgerald, 2006). In this context, some definitions are necessary. OSS 2.0 is defined as "... the more mainstream and commercially viable form ..." of F/OSS (Fitzgerald, 2006) or, as Conlon (2011) summaries, "[...] software designed to automate businesses of a particular type.

Table 25. The Three Primary Forms of Motivation

Identified Regulation	An individual is motivated to execute actions because they appear important and valuable
Introjected Regulation	When an individual takes pleasure in the outcome or senses guilt or shame
External Regulation	An individual is motivated to execute actions to receive external rewards or avoid punishment

Source: Li et al., 2012.

The Linux operating system was an excellent example of how roles were assigned by the participants without commercial organisations' involvement. But how were these individuals motivated? It is a question that has inspired many scholars in innovation management literature. According to self-determination theory (SDT), individuals have a two-fold motivation to take part in free and open-source software projects: either intrinsic or extrinsic (Deci and Ryan, 2008). Intrinsic motivation is driven by activities to experience satisfaction, either enjoyment-based, or obligation-based (morals, values and ethics). Extrinsic motivation is galvanised either by rewards such as career, prestige and positive evaluations, or to avoid punishment. Intrinsic and extrinsic motivation is divided into three primary forms: identified regulation, introjected regulation, and external regulation. The three types are explained in more detail in Table 25. The concepts of intrinsic and extrinsic motivation – the SDT theoretical lens – have been widely used in the IS(IS) field (Li et al., 2012).

A study conducted by von Krogh et al. (2012) on free and open-source software (Deci and Ryan, 1985; Gagné and Deci, 2005), explored different reasons that bring about human action intrinsically and extrinsically (Deci and Ryan, 1985) and several other empirical studies have demonstrated that free and open-source software developers are intrinsically and extrinsically motivated (Hars and Ou, 2002; Lakhani and Wolf, 2005; Roberts et al., 2006; Wu et al., 2007). Some scholars (Lindenberg, 2001; Lakhani and Wolf, 2005); Osterloh and Rota, 2007) have discussed it as enjoyment-based intrinsic motivation and obligation-based or community-based intrinsic motivation. Wu et al. (2007) also used the SDT framework to describe the continued motivation to contribute to free and open-source software projects. Hars and Ou (2002) found that intrinsically motivated developers spend more time and effort in open-source projects than their counterparts – externally motivated. Lakhani and von Hippel (2003) associated the perception of ability and fun to the willingness to help others. Other scholars have also identified motivation based on reciprocity (Bergquist and Ljungberg, 2001; Wu et al., 2007); reciprocal helping behaviour (Lakhani and von Hippel, 2003); or status motivation (Roberts et al., 2006). These works contradict Lerner and Tirole's (2002) early contribution to the topic, which was purely based on extrinsic motivation. Based on the work (taxonomy) of Feller and Fitzgerald (2002), Bonaccorsi et al. (2006) differentiate between economic, social, and technological motivation. Hemetsberger (2004) discussed motivation as "self-interest" and "others-orientation".

2.5.5. Free and Open-Source Business Models

Zott and Amit (2010, p. 222) define business models as: "a template of how a firm conducts business, how it delivers value to stakeholders (e.g., the focal firms, customers and partners) and how it links factor and product markets". Table 26 summarises open-source business models based on Watson and Boudreau (2008).

Table 26. Open-Source Business Models

Proprietary and Open Communities	The adoption of free and open-source technologies has increased significantly over the last few decades. Based on the level of adoption, it can be said that the quality of technology is more than satisfactory. Some firms have recognised an opportunity that arose from the level of acceptance and the technological gap in firms. As a result, many firms have emerged to fill the gap – Red Hat, Canonical etc. These types of firms have improved distributions methods and have developed complimentary services to make free and open-source products /services more accessible to a broader market.
Sponsored Open-Source	As the interest in the technology increased over the years, corporations have started to sponsor some free and open-source software projects. Corporations initiated some of the projects as closed code. Still, to accelerate the development of the technology, it was eventually released under an open-source licence.
Second-Generation Open-Source	Second-generation open-source can be understood as a hybrid between a corporate distribution and sponsored free and open-source software. Similarly, to the corporate distribution model, the revenue is built upon complementary service around the technology. On the other side of the spectrum, like sponsored projects, firms provide the best part of the development resources needed to create and maintain the software. But it differs from most corporate distribution companies, because firms do not sell licences in most cases. Unlike sponsored projects, the firms take great responsibility in governing the projects. Because firms can dictate the directions of the projects, they are very intimate with the codebase. They can exploit the technology to provide higher- quality service than could potentially competing service providers.

Source: Watson and Boudreau, 2008.

Thus, a business model consists of intertwined elements that characterise how a product/service can create value: the value proposition, the target customers segments, how the product/service are delivered, and the relationship is maintained (customer interface), which capacities, resources and network an organisation needs (infrastructure management) and how the products/service offered produce costs and generate revenues (financial aspects) (Osterwalder et al., 2005; Al-Debei, 2010). Business models in high-velocity markets rely heavily upon information and communication technologies (ICTs) as a facilitator and trigger of innovation (Nambisan, 2013). ICT is an appropriated unit of analysis in IS research (Rai and Tang, 2014). Veit et al. (2014) have pinpointed three pillars of business model research in IS: business models in IT industries, digital business models and IT support for developing and managing business models. Conceptual and theoretical contributions characterising and depicting the concept of a business model have dedicated the research in business models (Osterwalder and Pigneur, 2009) and linking it to associated fields such as strategy innovation and management (Porter, 2001; Chesbrough, 2010; Teece, 2010).

Organisations are regularly readjusting to turbulent environments and are often modifying their business models to stay successful (Wirtz et al., 2010). But, the literature on business models has so far neglected elements such as how and why firms change business models over time (Kijl and Boersma, 2010). Most recently, the gap has been recognised by scholars in the discipline of management, and the interest in a more dynamic perspective of business models has increased (Burkhart et al., 2011). It can now be said that research has identified some organisational capabilities of significant importance for firms to react to cycles of innovations (Demil and Lecocq, 2010; Aspara et al., 2011). The ability of firms to change their business models during phases of turbulence rests on the firm-specific environmental pressure to adapt (Linder and Cantrell, 2000; Madjdi and Hüsig, 2011) and can vary from minor to significant adaptions and revisions or the development of new or combined business models (O'Reilly and Tushman, 2008; Cavalcante et al., 2011). There have been some attempts to identify what is known about the business model change in the software industry. The works have shown that the majority of studies on the subject focus on emerging technologies such as cloud computing, development paradigms, open-source software and their possible effects on currently known business models (Kranz et al., 2016).

Open-source software includes all three of the elements that drive creative destructiveness. Most recently, numerous innovations have been built upon free and open-source software technologies. The zero-cost licence structure is one of the significant elements of the broader acceptance of the technologies even in previously untapped markets. Another critical component of the distribution of free and open-source software has been the internet. Free and open-source software relies on new ways of software development. It uses the public as the primary labour force and enables the management and location-agnostic access to technology (Watson and Boudreau, 2008). The authors have identified five models of software production or distribution. The five models are summarised in Table 23.

2.5.6. Ambidexterity in Free and Open Software

Open-source ecosystems are networks of different stakeholders, actors and entities, and are understood in the innovation literature as a basic form of OI networks. The networks assemble naturally and consist of “informal and fluid structures” that traditionally lack hierarchical governance (Fleming and Waguespack, 2007; O’Mahony and Ferraro, 2007). Open-source ecosystems embody the complexities of the complementary and opposing relationships among stakeholders (O’Mahony and Ferraro, 2007). The literature on inter-organisational relations debates the determination of actors to share knowledge in networks as exploitation and exploration. Parmigiani and Riversa-Santos (2011) discuss the two pure forms of inter-organisational relationships as “co-exploitation and co-exploration”. Yang et al. (2011) make a distinction between the learning in exploitation alliances versus exploration alliances: exploration alliances are associated to intrinsic value from tacit, new knowledge, and exploitation is linked to alliances with short-term benefits from existing knowledge. The two types of alliances offer different advantages to firms. At the same time, exploration alliances seem to be more dynamic and generate more opportunities by understanding the value of the partner's resources more accurately. Exploration orientation is usually smaller in size, consisting of less established organisations or new ventures. These types of firms aim to profit from getting entry to the assets of larger, more established actors to advance the exploitative capabilities needed for securing value from innovation. At the other side of the spectrum, more established "bigger firms" are very likely to exploit actors that explore more radical, disruptive innovations (Buckley and Prashantham, 2016) (in Faridian and Neubaum, 2018).

The coordination of software at the firm-level includes conflicting decisions between options such as efficiency *vs* flexibility (Adler et al., 1999), exploitation *vs* exploration (March, 1991), and evolutionary *vs* revolutionary change (Tushman and O'Reilly, 1996). Software project managers must maximise the performance of the current production processes while at the same time adapting them to specific customers' needs; software project managers must stress the repeatability of the development processes on the one hand and the ability to respond to dynamic business demands on the other (Boehm, 2002; Napier et al, 2006). Software project managers must take advantage of the current capabilities of the firm's relation to existing clients and at the same time pursue emerging technologies and business opportunities; they must ensure that product and project portfolios please existing customers while at the same time preparing for market growth (Markowitz, 1952; McFarlan, 1981; De Reyck et al, 2005). Additionally, software project managers need to invest in both evolutionary and revolutionary changes (Tushman and O'Reilly, 1996) such as incremental innovation (Humphrey, 1989; CMMI Product Team, 2006) and new project management techniques (Fonstad and Robertson, 2006; Woolridge et al., 2007). According to Lyytinen and

Rose (2003/2006), firms also need to consider radical innovations like open source and globally distributed software development (Napier et al., 2011).

2.6. SERVICE DIGITAL PLATFORMS

2.6.1. Service Science

As service science becomes increasingly evident as a scientific discipline, it calls for its own conceptual framework. The definition of service science consists of recommendations to studies of service systems. According to Spohrer et al. (2007), service systems are characterised by the value-coproduction configuration of people, technology, other internal and external service systems, and shared information. Wladawsky-Berger outlines that service systems involve ‘‘market-facing complex systems’’, which consequently are directly associated with economic exchange. Thus, there is the need for a conceptual foundation for service science grounded on market exchange processes. The processes must include people, interaction, innovation, learning and technologies (Lusch et al., 2008).

Maglio and Spohrer (2008, p. 18) define service science as an interdisciplinary field that "combines organisation and human understanding with business and technological understanding to categorise and explain the many types of service systems that exist, as well as how service systems interact and evolve to co-create value", and service systems as: "value co-creation configurations of people, technology, value propositions connecting internal and external service systems and shared information" (p. 18). Therefore, service systems are contemplated as the basic unit of analysis in service science. The network structures form the concept of "open systems" that can improve other systems by sharing and applying resources, or the ability to develop its own system by gaining external resources. Service systems formulate an abstraction able to be analysed within very different disciplines and industries (Spohrer et al., 2008). The size of the network structures can range in dimension from a single person to a global exchange system (in Vargo and Lusch, 2010).

2.6.2. The Goods and Services Debate

The intellectual discussion regarding what goods and services consist of can be traced back to Adam Smith. Classical economists such as Smith thought about goods (initially "commodities") as something over which ownership rights could be instituted and exchanged. The seminal *Wealth of Nations* (Smith, 1776/1969) differentiates the result of productive labour, which could be stored in inventories and subsequently exchanged for other commodities of value, from "unproductive" labour. According to Smith (1776/1969), "unproductive" labour is useful and necessary, but does not generate something of equal value that can later be exchanged. The early debates on the subject happened in a philosophical context that accentuated the significance of capital and capital creation. In the 18th century, the ownership of goods was associated with wealth. But calling the output of some members of the society[1] "unproductive" could maybe be seen as somewhat provocative. Nevertheless, it was only around the beginning of the 19th century that Jean Baptiste Say (1803/1964), a liberal French economist, coined the term 'immatériel' (immaterial or intangible). The economist uses the example of physicians who would patent something new but would not generate a tangible good or a product. Yet, physicians would exchange this patent for a fee. Say recognised the act of giving as the physicians' act of production. Although over the 19th century there was considerable discussion around the terminology used by Say, the term 'immatériel' has grown into a well-accepted definition of what differentiates goods[2] from services (in Lovelock and Gummesson, 2004).

1 Members such as: government officials, the armed forces, clergy, lawyers, physicians, "men of letters," musicians, entertainers, or "menial servants

2 The term goods was eventually implemented by Marshall (1890/1962) to replace commodities.

The contemporary concept of 'services' was coined by the U.S. Department of Commerce's Standard Industrial Classification (SIC) codes in the 1930s. The economy of the 1930s was dominated by the primary and secondary sectors, and the notion of 'services' was not understood in the same way that we understand it today (Chesbrough and Spohrer, 2006). Services were usually portrayed as "aids to the production and marketing of goods" (Converse, 1930; Fisk et al., 1993). By the 1950s and 1960s, every sector of the economy that could not be linked to manufacturing or agriculture was commonly referred to as the ''service sector''. The sector grew substantially to become more expressive in the number of jobs than the manufacturing sector (Clark1957). By the end of the 1950s, economists and politicians alike already had a much better understanding of how economic growth worked in the service sector. Until 1980, manufacturing dictated overall economic growth (Cohen and Zysman1988). But, with the increased expansion of ICT, the service sector gained progressively in relevance. The increased productivity in retail due to the introduction of bar code scanning, megastores, e-commerce, in addition to the comprehensive spread and introduction of financial and communication services, were critical for the growth of the service sector throughout the 1990s (Gadrey and Gallouj, 2002). According to Tien and Berg, (2006), most developed economies have 70–80% of their GDP and employment in the service sector, 15–25% in the manufacturing sector, and around 5% in the agricultural sector (in Maglio et al., 2009).

2.6.3. Service-Dominant Logic

As discussed in the last section, the classical conception of "productivity" has had a significant impact on the understanding of economics since the 18th century. The intellectual agenda was developed around a logic of tangible goods and productivity and has been ever since associated with the production of tangible goods that can be exported to enhance national wealth (Smith, 1776/1969). This agenda reflects the good-centric terminologies in commerce today: the expressions go from product to production, goods, supplier, supply chain, value-added, distribution, producer and consumer. These are certainly more than words, and express the underlying paradigm thinking for exchange in general. The long-standing tradition in thinking about economics poses problems for any effort to characterise and describe a counter-paradigmatic view, S-D logic. The main issue here is that there are not even terms available for a paradigm change. It is, therefore, preferred and generally accepted to use G-D logic terminologies to describe a new paradigm, SD logic (Vargo and Lusch, 2008).

Historically, services are affiliated with a very different logical category and economic philosophy. It can be said that there is a significant transition from a goods-based to a service-centric economy (Spohrer and Maglio, 2008; Maglio and Spender, 2008). Practitioners and academics share the opinion that firms should shift resources to become more service-oriented (Davies et al., 2007; Gebauer and Fleisch, 2007). The paradigm change is reflected in business-to-business (IBM and GE), business-to-consumer (Kodak, Apple) and software-as-a-service (Vargo and Lusch, 2008). Today, there are two dominant paradigms within the service system literature, goods-dominant (G-D) logic and service-dominant (S-D) logic. The goods-dominant logic views goods or 'tangible output' as a primary criterion for economic trade; and 'services' (plural) as some intangible goods or an add-on that embellishes the market price (Vargo and Lusch, 2004). The G-D logic is also discussed in economics as "neoclassical economics research tradition" (Hunt, 2000), "manufacturing logic" (Normann, 2001), "old enterprise logic" (Zuboff and Maxmin, 2002), "marketing management" (Webster, 1992), and an "output pulling paradigm" (Zhao et al., 2007). The G-D logic is organised around goods, and focuses on 'products': tangible (goods) and intangible (services), as the units of output. The essence of G-D logic is the transaction of products as the primary unit of output, and its values is added during the process of manufacturing, farming or extraction. Ideally, the production occurs separately from the client, and outputs are standardised (Vargo and Lusch, 2008).

The other logic (S-D logic) considers ‘service’ (singular) as the activity of executing something for someone, with no association to goods. It acknowledges ‘service’ as the foundation of exchange, the service-dominant (S-D) logic (Vargo and Lusch, 2004). The main distinction between G-D and S-D logics is the approach to service. While S-D logic considers the use of abilities (knowledge and skills) for the benefit of others and the word ‘service’ is in the singular, G-D logic deliberately adopts the plural ‘services’. Additionally, it embodies a change from anticipating value in terms of “operand resources” (tangible resources that necessitate some activity to add value to them) to “operant resources” (intangible resources that can create value) (Madhavaram and Hunt, 2008).

Most importantly, G-D logic views services as inferior to goods, whereas S-D logic thinks of service as a process of accomplishing something for another party. Consequently, the creation of value is seen as a collaborative process rather than from a ‘producer’ perspective. S-D logic is not about a different object being investigated but an actual shift in logic to a process-driven, service-centric logic (Vargo and Lusch, 2008).

2.6.4. Service Innovation

Service consists of applying resources for the benefits of others or oneself. It is line with operations in businesses, government, non-profits, households, and individuals (Lusch and Nambisan, 2015) using service models from IT, such as “service-oriented architecture”, “software-as-a-service”, and, more generally, service computing (Zhao et al., 2007). Large organisations across industries have acknowledged service as a means for business development. Businesses are going through a fundamental shift due to new requirements in service and product development (Lyytinen and Rose, 2003). Organisation complexity, combined with technological innovation and the increased demand for agility, has transformed the way firms and academia think about IS(Dermikan and Goul, 2006). Organisations have to undergo structural changes, both operational and cultural, to become more service-oriented. According to Brown et al. (2006), organisations need to rethink their "traditional siloed and tightly coupled" ways of doing business and develop "loosely coupled services" and align them vertically with IT services that are sourced by virtual resources.

Changes involving complex information and communication technologies (ICTs) need time. Products and procedures mature over time and are path-dependent (Clark, 1985). Innovation takes place by "recombining or rewrapping" resources, and points out that the more significant the investment in innovation the more prominent the potential for products/service being introduced to different markets. Each innovation is a new building block that can be "recombined or rewrapped" with other resources and potentially become a new innovation (Lusch and Nambisan, 2015). Service computers are excellent enabler for business agility which is said to improve productivity, shorten development cycle times and deliver flexibility (Demirkan et al., 2008). According to Barras (1990), ICTs have been understood as technological tools in the service delivery process, which contribute to productivity and efficiency of service firms and which may, over time, lead to entirely new markets or categories of services. It boils down to how organisations apply their IT assets. As Sambamurthy et al. (2003, p. 244) have observed, "IT competence is the organisational base of IT resources and capabilities and describes a firm's capacity for IT-based innovation by the available IT resources and the ability to convert IT assets into strategic applications".

One key element of information technologies (IT) competence involves: (1) the number of IT investment; (2) the excellence of IT in the organisation; (3) the IT human resources; (4) the type of IS/business conglomerate (Feeny and Wilcocks, 1998; Henderson, 1990; Ross et al., 1996; Weill and Broadbent, 1998). When ICTs are combined with other internal and external resources in firms, ICTs facilitate the transfer of knowledge that can be rearranged in different environments to create new openings for service interchangeability and innovation (Lusch and Vargo, 2014). Current efforts to understand digital infrastructure (Tillson et al., 2010) have stressed the very diverse characteristic of digital technologies (Henfridsson and Bygstad, 2013), which may accelerate service innovation (Yoo et al., 2012). According to Den Hertog (2000), there are four dimensions of novelty of service innovation: (1) service concept; (2) client interface; (3) service delivery system; and (4) technology, with some service innovations linking the specific arrangement of these four dimensions.

Service is delivered to specific customers in very particular contexts and should be examined as an evolving, interactive, dynamic process that involves the transfer of knowledge and information, with communication channels between providers and customers (Miles, 2008). The convergence of products and service is evident in the emergent tendency toward servitisation, and it is in line with the development of the IT field (Neely, 2008; Vandermerwe and Rada, 1988). The literature of IS(IS) regards digital artefacts as dominant over digital innovation; hence, the importance of digital artefacts for service innovation. Digital artefacts are characterised as enjoying an uncertain logic (Kallinikos et al., 2013), being consciously imperfect and repeatedly re-establishing themselves (Garud et al., 2008; Zittrain, 2008). Digital artefacts have been interpreted similarly, but still in very different ways, in the IS(IS) literature (Faulkner and Runde, 2009; Kallinikos et al., 2013; Yoo et al., 2010). According to Kallinikos et al. (2013, p. 357), “[…] as digital artefacts become increasingly embedded in wider and constantly shifting ecosystems, they become editable, interactive, reprogrammable, and distributed."

2.6.5. Service Ecosystem

Hannan and Freeman's (1977/1989) seminal work on organisational ecology has initiated the wide-spread endorsement of an ecological metaphor in business and organisational study. Academia and practitioners have recognised the value of the ecosystem framework to understand the complex structures of business relationships (Harte et al., 2001). The term 'technology ecosystem' emphasises the organic surroundings of technological progress and innovation that is often incomplete in more traditional approaches. The view of an 'ecosystem' in biology refers to a diversity of living organisms that are in line, influenced and changed by diverse external dynamisms (Adomavicius et al., 2008). According to Adomavicius et al. (2007), biological ecosystems consist of a population of organisms (including enemies such as predators, parasites etc.), resources and external forces. The concept of an ecosystem represents a way of describing a set of coevolving technologies.

Successful businesses evolve, and innovation does not mature in a vacuum. These types of companies need resources, capital, partners, suppliers, and customers to co-create cooperative networks (Moore, 1993). An ecosystem is a combination of interconnected bodies (organisations and individuals) that use their competence and activities and build upon one another for existence. The approach is critical to demonstrate the diverse technologies and relationships that structure IT settings (Adomavicius et al., 2008). Lusch and Nambisan, using the concept of S-D logic, argued that service innovation is embedded in an actor-to-actor (A2A) network and has its roots in the idea of service ecosystems, which underline the significance of organisational structures and the arrays of principles to enable resources integration and service exchange among stakeholders.

The concept of service ecosystems is based upon the idea that ICTs enable service innovation. Digital platforms stand for the underlining theories upon which firms can develop complementary goods, technologies, and services (Gawer, 2009) in addition to digital competences (Yoo et al., 2012). The "ecosystem model of technology evolution" (Adomavicius et al., 2007) incorporates various design and theoretical approaches in economics, engineering, and organisational theory.

Service ecosystems are self-governing, self-adaptable systems of loosely coupled social and economic environments, in which stakeholders communicate with each other through a shared infrastructure and co-create value jointly through service exchange (Lusch and Nambisan, 2015). Software and regulations, the so-called boundary resources, are used as an interface for the dialogue between the platform owner and the developers of applications. The setting grants innovators the connection to the core resources of service systems. It encourages generativity, while simultaneously enabling firms that developed the infrastructure some control over the digital ecosystem (Ghazawneh and Henfridsson, 2013). Lusch and Nambisan (2015) have pointed out three critical concerns related to the service ecosystem under the S-D logic: (1) the need to reinforce both the structural flexibility and the structural integrity of the service ecosystem and administrate the conflicts between the two; (2) the necessity to advance and manage shared beliefs (philosophies) among stakeholders; (3) the need to formulate and put in operation an architecture of stakeholders to delegate them and exchange service.

The types of service just discussed are akin to re-useable objects that consist of repeatable deeds and duties, and can enter through arrangements, enabling organisations to deal with changes in the environment through agility (Peltz, 2003; Wilkes, 2006). Agility can be understood as the ability to recognise favourable circumstances for innovation and allocate the opportunities by mobilising resources, insights, and affiliations with rapidity and flexibility (D'Aveni, 1994; Goldman et al., 1995). Agility consists of the exploration and exploitation of market potential arbitrage. While exploration is the managerial trial and error and searches for knowledge on unspecified changes for competitive activities, exploitation is the exercise and improvement of products already known and the enhancement and expansion of current abilities, technologies, and know-how (March, 1991). To adapt to the requirements of the various actors and to exploit the indirect network effects, platform owners stimulate innovators outside their organisational boundaries to co-develop the platform (Linder et al., 2003). According to Gawer and Cusumano (2002), the model of platform ecosystems has its roots in the concept of complementary innovation: platforms become increasingly valuable, the more the network of innovation expands. There is a stream of literature trying to define how ecosystems are formed, and the consequences for platform owners, third-party suppliers, and consumers (Adomavicius et al., 2007/2008; Eisenmann et al., 2009; Gawer and Henderson, 2007; Lee and Mendelson, 2008; Mantena et al., 2007; Parker and Van Alstyne, 2008; West, 2003); other scholars (Eisenmann et al., 2009; Parker and Van Alstyne 2008; West 2003) explored how platform owners can motivate the advancement of ecosystems.

Platforms ecosystems vitalise an environment for an investigation into the dynamic pattern of inter-organisational collaboration and competition (Ceccagnoli et al., 2012). Some authors (Eisenmann et al., 2009; Parker and Van Alstyne, 2008; West, 2003) have considered organisational strategies in platform marketplaces, which has included some evaluations on the degree of openness of a platform, while others (Lee and Mendelson, 2008; Mantena et al., 2007) discussed the retailers' feedback to open and closed platforms. Digital platforms are always going through updates and changes; hence, the instability of digital platforms. Digital artefacts are embedded into networks of technical and structural interactions that highlight their uncertainty and ambivalent logic. The types of digital artefacts imply, and gradually further, the development of other digital artefacts within the platform and are often dispersed throughout different contexts. This arrangement can be understood as the building blocks of even larger digital ecosystems. It is imperative to note that these kinds of ecosystems never rest. The artefacts they are built upon undergo constant changes, as the interactions to which they are exposed fluctuate (Kallinikos et al., 2013).

The boundaries are applications and guidelines that perform the function of a channel for the interactions between the platform owner and the application designer. Platform boundary resources are critical for third-party application designers (von Hippel and Katz, 2002), who are seldom directly rewarded for their design work. Instead, designers are given access to particular markets for their designs, which potentially offers them a superior capacity for service innovation (West and Mace, 2010). When seen from the perspective of the platform owner, they benefit from the differentiation of their products through expansion (Boudreau, 2011). Additionally, platform owners enter into dialogue with various networks of designers, characterised by unrelated innovation skills and knowledge resources (Boland et al., 2007; Yoo et al., 2010).

To be able to develop a concept about the idea of information technology (IT) ambidexterity, it is critical to detail all components of information technology (IT) capabilities. According to Bharadwaj (2000), these capabilities are the ability to "mobilise and deploy IT-based resources in combination or co-present with other resources and capabilities" (Bharadwaj, 2000, p. 171). Categorically, it is the firm's information technology (IT) capabilities which consist of tangible information technology (IT) assets such as the settings that establish the information technology (IT) infrastructure capabilities (Ajamieh et al., 2000). Information technology (IT) competencies are an important part of information technology (IT) capabilities as they enable intra- and inter-organisational business processes, as well as knowledge management through the application of information technology capabilities (Zhu, 2004). To capitalise on information technology's (IT) infrastructure, e-business organisations need to "explore and exploit" their information technology (IT) capabilities. The processes of "explore and exploit" are in line with Levinthal and March's (1993) ideas of how firms seek either exploitation for efficiency or exploration for innovation, or both at the same time. According to Chi et al. (2017), information technology (IT) ambidexterity can be characterised as the firm's capacity for applying information technology (IT) capabilities for both exploration and exploitation to improve performance (in Ortiz de Guinea and Raymond, 2019).

The interconnectivity of dispersed actors and the associated digital resources in service ecosystems is a great service innovation facilitator. In digital ecosystems, organisations need to exploit and explore innovation at the same time; hence, ambidexterity. Digital transformation outlines an induced technology change that consists of the exploitation of digital technologies to enhance existing processes and the exploration of digital innovation (Berghaus and Back, 2016). The expression "digital transformation" is widely used in the field by scholars and practitioners alike (Morakanyane et al., 2017). Often digital transformation is used with digital technologies, meaning devices that are capable of process binary information in a digital network. The devices are interconnected, providing a platform where actors can interact with each other and share information (Hadlington and Scase, 2018), resulting in innovative services (Ryu and Lee, 2018). According to Vargo and Lusch (2004), service is designed via collaborative processes in which knowledge and skills (operant) are combined with tangible resources (operand) to create value to stakeholders in service systems. Service innovation can be defined as a fundamentally new process or service, offering an addition or change in the delivery process. The complexity combined with rapid changes in markets and technologies (D'Adderio, 2014) is said to have a significant impact on competition. Organisations require to speed up cycle times of exploring new service innovation and exploiting them methodically (Ojasalo and Ojasalo, 2018). The management of the two capacities, exploration and exploitation, is in line with concepts of ambidexterity (Röglinger and Schwindenhammer, 2018). Exploration is related to radical service innovation to advance changes in service ecosystems (Röglinger and Schwindenhammer, 2018). It, in turn, can put organisations at risk because services take time to be adopted by clients (Witell et al., 2016). In addition to exploring radical new service innovation, organisations have to exploit existing services through continuous improvement to gain efficiency (Röglinger and Schwindenhammer, 2018). Organisations need to adapt to changing environments (O'Reilly and Tushman, 2013). Consequently, they are exposed to conflicting requirements: to achieve innovation through the exploration of new competencies and to improve efficiency by exploiting their current capabilities (Duncan, 1976).

2.7. CONCLUSION

This chapter has described the key theoretical concept of the thesis. As a starting point, we explored different definitions of innovation by various scholars in the management science literature. Then, we proceeded to typologies of innovation and evolving models of innovation. The study has demonstrated that innovation is an evolutionary and interactive process. Additionally, the chapter presented the literature review of inbound and outbound OI, free and open-source software and service science. The literature review has shown that outbound OI is less explored in theory and practice and it is from this that the research gaps were identified. These are: (1) the lack of evidence about outbound OI processes of technology consumers and creators in firms' current markets; (2) the lack of works on outbound OI processes in free and open-source software service digital platforms ecosystems, and (3) the thesis answers to recent calls for more research on ambidexterity in outbound OI processes in firms.

In combination with the research objectives of the thesis, the research gaps informed the two research questions presented in Section 1.2, Research Questions. Afterwards, this thesis discusses two frames of literature critical to the understanding of the demonstrator of the thesis (free and open-source software and digital service platforms – the Apache Hadoop). The Apache Hadoop is covered later in the thesis in Chapter 5, The Apache Hadoop.

CHAPTER 3. THEORETICAL FRAMEWORK

3.1. INTRODUCTION

The previous chapter analysed the OI and the open-source literature based on which the research gap was established. The research gap was central to the development of the two research questions. The literature review also introduced the concept of service digital platform as a complementary stream of literature to prepare the reader rationally to the demonstrator of this work in chapter 5, the Apache Hadoop.

This chapter introduces the theoretical framework of the study not only based on the literature review in the previous section but also inspired by the demonstrator itself. There are two critical aspects of innovation streams that is of major importance to understand how innovation takes places and evolves within the Apache Hadoop ecosystem. One of which is the concept of ambidexterity (Tushman et al., 2010). Organisational ambidexterity deals with the dynamic capabilities of firms to explore and exploit, that is, to compete in established technologies and markets where expertise, management, and incremental change are valued and where emerging technologies and markets require flexibility, autonomy and experimentation. Although various frameworks have been employed to describe organisational ambidexterity, this work applies ambidexterity from a dynamic capabilities perspective. For example, Teece et al. (1997, p. 516) describe dynamic capabilities as "the firm's ability to integrate, build, and reconfigure internal and external competencies to address rapidly changing environments". Similarly, Helfat et al. (2007, p. 1) characterise dynamic capabilities as "the capacity of an organisation to purposefully create, extend, or modify its resource base" (in O'Reilly and Tushman, 2013).

Apache Hadoop creators and consumers explore and exploit their dynamic capabilities to develop processes to externalise their knowledge and commercialise their technologies in their current markets. Thus, innovation streams are used in this work from a technological change perspective (the ability of firms to explore and exploit their technical capabilities) and not necessarily from an organisation design angle. An innovation stream is a portfolio of products and services managed simultaneously by an organisation or strategic business unit (Tushman & Smith, 2002). Tushman and O'Reilly (1996) and Tushman & Smith (2002) understand that the notion of innovation streams might be used as tools to illustrate how technology may evolve. The portfolio of products and services is defined as the technology and target markets of the firms (Abernathy and Clark, 1985). The firm's innovation relating to products and services is defined as incremental (Christensen, 1997; Dosi, 1982), architectural (Henderson and Clark, 1991) or discontinuous (Gatignon et al., 2002).

Secondly, there are different streams of research, trying to understand innovation and technological change. These studies are of significant importance for the understanding of the technological cycle. Dewar and Dutton (1986), Ettlie et al. (1984), and Damanpour (1996) have discussed these cycles as "discontinuous or radical" vs "incremental". Tushman and Anderson (1986/1990) have coined the terms of "competence-enhancing" vs "competence-destroying". While Henderson and Clark (1990) explore "architectural" and "generational", Christensen and Rosenbloom (1995) formulated "disruptive" innovation. Clark (1985) and Tushman and Murmann (1998) developed the concepts of core/peripheral. Baldwin and Clark (2000) and Schilling (2000) advanced the notion of "modular". Table 27 summarises these different streams of researches.

Table 27. Different Streams of Research

Technological Cycles	
Authors	**Definition**
Dewar and Dutton (1986), Ettlie et al. (1984), and Damanpour (1996	"discontinuous or radical" vs "incremental"
Tushman and Anderson (1986/1990)	"competence-enhancing" vs "competence-destroying"
Henderson and Clark (1990)	"architectural" and "generational"
Christensen and Rosenbloom (1995)	"disruptive"
Clark (1985) and Tushman and Murmann (1998)	core/peripheral
Baldwin and Clark (2000) and Schilling (2000)	"modular".

Consequently, firms have to completely rethink the way they approach product and service design (Abernathy and Clark, 1985). At the core of organisational adaptation is a firm's ability to exploit its current capabilities as well explore into future opportunities (March, 1991; Levinthal and March, 1993). One manifestation of a firm's ability to explore and exploit is its ability to initiate innovation streams (Katila and Ahuja, 2002; Tushman and Smith, 2002; Tripsas, 2009). Innovation streams are portfolios of innovation that include incremental innovations in existing products as well as substantial innovation that extend the existing technological trajectory and/or move it into different markets (Abernathy and Clark, 1985; Eisenhardt and Tabrizi, 1995; Venkatrman and Lee, 2004).

The second important aspect of innovation streams is that, although the studies mentioned above are of immeasurable value for the field, they perceive innovation as a binary system or system consisting of two classes (e.g. "discontinuous or radical" vs "incremental"; "competence-enhancing" vs "competence-destroying"; "architectural" vs "generational"; "core vs "peripheral"), taking for granted the differences of innovations when they do not follow a logical continuity. Unlike these binary concepts introduced above, the lens of innovation streams offers one additional parameter for the understanding of technological cycles within the Apache Hadoop. This further classification of innovation presents a definite advantage to understanding a platform (Apache Hadoop) that consists of many different modules, architectures and markets.

3.2. OVERVIEW, OBJECT AND UNIT OF ANALYSIS

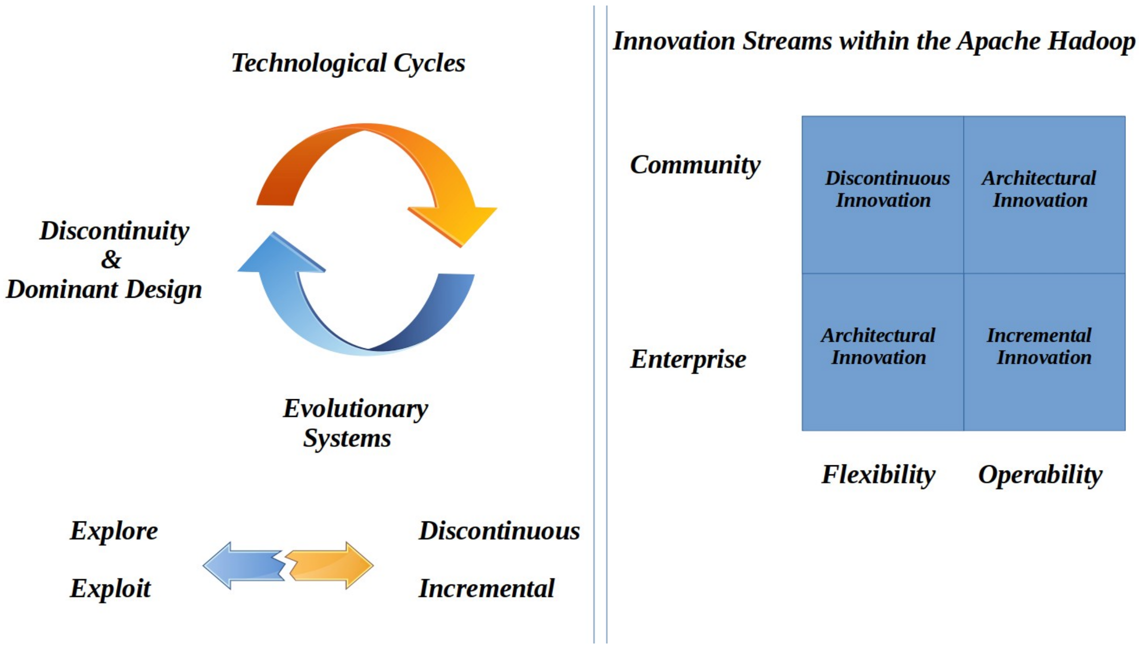
3.2.1. ***Overview***

Figure 2 demonstrates the theoretical framework of the study. It aims to provide an overview of how innovation streams operate in the outbound OI processes of technology creators and consumers in their current markets. On the left-hand side, the preconditions of innovation streams are presented. These consist of evolutionary systems (from discontinuity to a dominant design). It reflects the dynamic capabilities of firms to explore new horizons (discontinuous innovation) and exploit current knowledge (incremental innovation). On the right-hand side, the two-by-two matrix portrays two distributions (community and enterprise) and four possible conditions depending on the flexibility and operability of the Apache Hadoop.

The top left corner illustrates a flexible community distribution that is critical for discontinuous innovation. The top left portrays a community distribution that has become more operational over the years but still offers some flexibility. The same can be said of the enterprise distribution at the bottom left. Both distributions share some flexibility and some operability, and they can be very well combined with other modules of the Apache Hadoop because the core of both distributions is precisely the same. Finally, the bottom right demonstrates the operational and incrementally evolving enterprise distribution.

As it will be discussed in this chapter, the theoretical framework focuses on the four possible conditions in the evolutionary process of the Apache Hadoop from the community base to enterprise distribution. In doing so, the study considers the generation of innovation streams and how the Apache Hadoop evolves from community distribution to enterprise distribution with a service-level agreement.

Figure 2. *Theoretical Framework*



3.2.1. Object of Analysis

IS research frequently borrows concepts from other scientific fields to describe or reflect on the data. IS researchers alter or combine current theories with building new theoretical frameworks that fit better to different intellectual needs. Firms across industries have changed their formerly hierarchical, linear supply chains into flexible networks of strategic relationships with external stakeholders (Bitran et al., 2007). As technological development and expert knowledge become more dispersed, businesses are shifting their innovation focus away from their internal research and development facilities and toward outside their firm's borders, enabling collaborative innovation and R&D (Ritala et al., 2013; Baldwin and Hippel, 2011). This reflects the acceleration of the trend toward more linked and collaborative corporate processes by digitally enabled networks, which open up new possibilities for less predetermined and more distributed organisational processes (Pagani, 2013). According to Ritala et al. (2013), the notion of ecosystems has recently acquired momentum among scholars, practitioners, and policymakers as it enables them to examine the interdependence and interactions of diverse players (Guggenberger et al., 2020).

The work highlights the types of interdependence and interactions in the context of outbound OI in digital innovation ecosystems and introduces innovation streams as a framework for examining how digital innovation evolves and diffuses. The object of analysis in innovation studies can be within organisations, across firm networks, at the macro-level or within communities (Garud et al., 2013). Thus, this work endorses macro-level analysis for digital innovation while referring to a complex composition of technology, techniques, concepts, commercial application, firms, and institutional settings as a "digital innovation ecosystem".

Both industry and academics have demonstrated a strong interest and focus on the benefits that big data can provide, and investment in big data skills and resources has risen in recent years (Gandomi and Haider, 2015; Goes, 2014; Gupta, et al., 2018; Tambe, 2014; Yaqoob et al., 2016). Big data is especially intriguing as a digital innovation subject because it is disruptive and transformational throughout the information value chain (Abbasi et al., 2016; de Camargo Fiorini et al., 2018). Thus, big data is an innovation that encompasses more than simply digital technology and tools; it includes knowledge, skills, concepts, organisations, and other social and institutional settings. In other words, the disruption does not exist in a vacuum. According to scholars (Chae, 2014; Targio Hashem et al., 2015; Sagiroglu and Sinanc, 2013), innovation in big data (like other innovations) has progressed throughout time (from business intelligence and data mining to data warehousing) and by incorporating new resources, e.g. analytical platforms (R, Python), computing architecture (high-performance computing), data processing frameworks (Apache Hadoop), infrastructure (cloud computing, data centres), analytical talents, beliefs, methodologies and regulations (in Chae, 2019). Thus, the object of inquiry in this study is a digital ecosystem, more specifically data processing frameworks (Apache Hadoop).

3.2.3. Unit of Analysis

To contextualise further and to connect the object with the unit of analysis, the importance of dynamic capabilities within the concepts of innovation streams and outbound OI must be emphasised. There is an extended literature (Ancona et al., 2001; Benner and Tushman, 2002; Dougherty, 1992; Eisenhardt and Martin, 2000; Levinthal and March, 1993; and March, 1991/1996/2006) discussing firms' capabilities to explore new knowledge and exploit existing knowledge. According to Raisch and Birkinshaw (2008), researchers have emphasised the relevance of balancing contradictory tensions (Adler et al., 1999; Brown and Duguid, 2001; Katila and Ahuja, 2002) and numerous studies have highlighted the necessity for businesses to balance exploitation and exploration activities, from technology change to organisational design (Benner and Tushman, 2003; Burgelman, 1991; Eisenhardt and Martin, 2000; Gupta et al., 2006; Tushman and O’Reilly, 1996; Volberda, 1996).

The thesis applies innovation streams from a technology change perspective to understand how creators and consumers of a digital ecosystem externelise their knowledge and commercialise their technology. Anderson and Tushman (1990) suggest that discontinuity is associated with a different approach to the design of products and services, and discontinuity can alter the underlying mechanism of the products and services themselves. The authors divide the technological cycle into four parts: technological discontinuities, eras of ferment, dominant designs and eras of incremental change. Dominant designs and technological discontinuities are events that mark the transitions between eras of ferment and eras of incremental change. Similarly to Anderson and Tushman (1990), this study divides the technological cycle of Apache Hadoop (the path from the community distribution to an enterprise distribution with compliance) into four parts. The concept of the technological cycle is discussed in more in section 3.5.1. Technological Cycles .

Dynamic capabilities are distinguished capabilities for assimilating, architecting and reorganising internal and external resources and competencies to address and possibly shape rapidly changing business environments (Teece et al., 1997; Teece, 2007). According to Teece (2012, p. 1396), “strong DC are critical to success, especially for pioneering a market or a new product category of an innovating firm. Dynamic capabilities consists of three clusters of activities: (1) identification and assessment of an opportunity (sensing), (2) mobilisation of resources to address an opportunity and to capture value from doing so (seizing), and (3) continued renewal to execute and sustain the innovation at scale (transforming).

These activities must be performed better if the firm is to remain relevant as markets and technologies change. Basically, outbound OI also depends on dynamic capabilities. According to Appleyard and Chesbrough (2017), having technological capabilities to support outbound OI is critical. Thus, the relationship between dynamic capabilities, outbound OI and ambidexterity is complementary (Chesbrough et al., 2018).

The firm's capability to innovate can be understood as the skills and know-how required effectively to absorb, master and improve existing technologies and create new ones (Lall, 1992). The capabilities of firms are critical in providing and maintaining a competitive advantage and in the design and implementation of a strategy (Grant, 1991). According to Eisenhardt and Martin (2000), ambidexterity mirrors circumstances under which dynamic capabilities are most helpful. O'Reilly and Tushman (2008) highlight the significance of ambidexterity as a dynamic capability and establish a parallel between dynamic capabilities and ambidexterity by referring to Teece's (2009) sensing, seizing, and reconfiguring. Consequently, dynamic capabilities can be understood as unique strategic and organisational activities (Eisenhardt and Martin, 2000). According to Helfat et al. (2007), dynamic capabilities are organisational procedures developed and implemented into firms to enable technological change and innovation. Guerra et al. (2016) suggest that being ambidextrous is a way for organisations to generate dynamic capability. O'Reilly and Tushman (2008) describe ambidexterity as a dynamic capability that improves a company's performance in changing environments (Popadiuk et al., 2018). In other words, when exploitation and exploration are understood as dynamic capabilities they are useful metrics for the better understanding of digital innovation ecosystems and their evolution.

Other authors have discussed dynamic capabilities in the paradigm of OI in the context of open-source software development. For instance, outbound OI raises questions about business model selection and technological strategy and can be associated with the advancement and management of such dynamic capabilities (Bogers et al., 2019). Building on this rationale, the unit of analysis takes the perspective of the firms' ability to exploit and explore their outbound OI processes for technological change and innovation. This study understands that outbound OI and ambidexterity are dynamic capabilities used by Apache Hadoop creators and consumers to externalise their knowledge and commercialise their technology. Dynamic capabilities are used as a metric to understand how firms create and maintain a competitive advantage in a continuously changing environment. The concept of dynamic capabilities (explore new knowledge and exploit existing knowledge), borrowed from the framework of innovation streams, will extend the current studies on outbound OI by describing and illustrating some conditions under which ambidexterity can help organisations to externalise their knowledge and commercialise their technology. This study emphasises the relevance of ambidexterity in outbound OI processes by focusing on firms' dynamic capabilities to exploit and explore their knowledge and technology. It pays special attention to how Apache Hadoop creators and consumers exploit and explore outbound OI processes in different industry sectors in Brazil.

The two research questions investigate the conditions under which outbound OI and ambidexterity (dynamic capabilities) can be mobilised to understand how Apache Hadoop creators and consumers externalise their knowledge and commercialise their technologies in the firms' current markets. Based on innovation streams (Tushman & Smith, 2002; Benner & Tushman, 2003; O'Reilly & Tushman, 2008; Tushman et al., 2010), this work examines how Apache Hadoop creators and consumers apply their outbound OI to exploit and explore their knowledge and technologies, and how these dynmic capabilities influence the evolution of the ecosystem over time. The objective of this work is to illustrate the transition from exploitation to exploration and vice versa (from discontinuity to a dominant design), and how these capabilities are critical for the evolutionary system of the Apache Hadoop. Thus, innovation streams offer a theoretical lens that helps to explain this.

3.3. PRECONDITIONS

The interest in the role of innovation in economic and social change has expanded, especially within the social sciences (Fagerberg, 2006). Innovation is accountable for improving cost-reduction strategies, introducing new products and services, and most importantly, developing innovative ways to coordinate business operations. Innovation has been influential in defining the society in which we live in today and the practices directly related to innovation, and technological change is identified as essential for firms (Lorenzi et al., 1912; Veblen, 1899; Schumpeter, 1934). Organisations need to evolve to respond to changing needs and preferences of clients to take advantages of favourable circumstances emerging from digital technologies and shifting marketplaces, systems and dynamics. There are different models, frameworks, classifications and definitions of innovations. The study applies an extended binary model or a two-dimensional perspective of innovation proposed in the 1970s and 1980s (Bantel and Jackson, 1989; Daft, 1978; Damanpour, 1999; Damanpour and Evan, 1984; Evan,1966).

The literature of innovation streams suggests evolving models of innovation. It provides views on the complex reliance on and the potential for various interactions between the different aspects of the innovation process (Edquist and Hommen, 1999). Accordingly, evolving models of innovation make it possible to determine the factors that shape and influence the innovation processes as it is built upon the theories of interactive learning (Lundvall, 1992) combined with approaches of the evolutionary theory of technical change (Saviotti and Metcalfe, 1991; Nelson and Winter, 1982).

As it will be discussed in more detail throughout this chapter, innovation streams are a firm's capacity to continue to leverage its existing resources (exploit) and to pursue future opportunities (explore) (March 1991; Levinthal and March, 1993; Katila and Ahuja, 2002; Tushman and Smith, 2002). Innovation streams are innovation portfolios that combine incremental innovation with at least one radical innovation (Abernathy and Clark, 1985; Eisenhardt and Tabrizi, 1995; Venkatrman and Lee, 2004). The literature of innovation streams recognises that firms which engage with ambidextrous design are more in the position of exploiting and exploring their products and service simultaneously. According to Tushman and O'Reilly (1997) and Tushman & Smith (2002), the notion of innovation streams can be used as a tool to illustrate how technology may evolve over time. Thus, the lens of innovation stream can be a useful instrument in this type of research and help to explain: (1) how Apache Hadoop creators and consumers generate innovation streams in their current markets (from discontinuous to architectural and incremental); And (2) how the Apache Hadoop can evolve from the community to the enterprise distribution.

Ambidexterity is regarded as one of the critical agents for a firm's long-term survival; the significance of ambidexterity is known to be of relevance to OI research and practice (Faems et al., 2005). Furthermore, in markets characterised by rapid technological change and innovation and short product life cycles – such as the Apache Hadoop – firms are reliant on novelty and efficiency (Alegre et al., 2006). The literature on OI advocates that firms should combine internal and external knowledge in their innovation processes and align the sources of innovation with the firms' business models (Cassiman and Veugelers, 2006; Chesbrough, 2006). Firms should work together with other actors to develop their innovativeness and preserve the firms' success (Ferraris and Santoro, 2014; Carayannis et al., 2015). Hence, the theoretical lens of innovation streams is an appropriated tool to understand the role played by the community and enterprise distributions in firms' outbound OI processes.

The work defends the opinion that OI is comprised of inter-organisational life cycles. Innovation is the product of the interactions between these different actors (Ferrary, 2011) – OI implies the ability to handle contrasting knowledge (internal and external) simultaneously. Thus, managing the exploitation – of internal knowledge – and exploration – of external knowledge – are essentially dispute rich (Lubatkin et al. 2006).

Due to the rapid technological change and internationalisation of markets, firms are no longer able to develop solely on their own – the have become ambidextrous instead – firms today are increasingly using both internal and external source of innovation (Chesbrough, 2003b; Chesbrough and Bogers, 2014). Consequently, the balance between radical/incremental innovation and internal/external sources of knowledge have gained on the importance of theory and practice (Ardito et al. 2018). Ambidexterity in the OI processes of firms is paradoxical from the externalisation of the firms' activities through outsourcing or by establishing alliances (Baden-Fuller and Volberda, 1997; Holmqvist, 2004; Lavie and Rosenkopf, 2006; Rothaermel and Deeds, 2004). Last but not least, ambidexterity reflects the degree of openness of the inbound OI processes in firms (Herzog and Leker, 2007). To summarise, OI consists of two different pillars: external/internal technology sourcing and internal/external technology commercialisation (Herzog and Leker, 2007), where knowledge flows through a semipermeable corporate membrane, leading to inbound and outbound OI (Chesbrough and Crowther, 2006).

To begin with, an organisation capable of harnessing its expertise by moving it beyond its border will generate value by cultivating or leveraging additional capital/knowledge (Teece, 1986). While building value banks on the possession of useful information (assets), it is highly dependent on the firms' ability to desorb the knowledge to leverage it successfully. Schulze et al. (2014) equate desorptive capacity with the coexistence of exploratory and exploitative learning. In other words, organisational ambidexterity is the ability of firms to explore and exploit knowledge (Tushman and O'Reilly, 1996; Rothaermel and Alexandre, 2009; Tamayo Torres et al., 2014/2017; Wei et al., 2014).

Finally, outbound OI helps firms to manage their technical resources properly. When firms disclose their innovation, they potentially gain market approval of their internal capacities (Kauppila, 2010). By defining the variations and reciprocal characteristic between internal exploration and external exploitation, companies may gain useful expertise in finding and maintaining comparable differences and complementary features in their exploration and exploitation activities. Exploitation and exploration allow businesses to interact with their ecosystem – such as the Apache Hadoop – and create synergistic benefits from distinct expertise linked to exploration and exploitation while promoting ambidextrous behaviour (Chang and Shiu, 2019).

On the one hand, the Apache Hadoop is open-source software. Open-source communities are networks of multiple users which are recognised as a fundamental type of OI networks in the literature on software. The networks typically combine and consists of "informal and complex systems" historically lacking in structured government (Fleming and Waguespack, 2007; O'Mahony and Ferraro, 2007). Open-source communities reflect the dynamic of stakeholder partnership, both complementary and competing (O'Mahony and Ferraro, 2007). The literature on inter-organisational relations discusses the perseverance to share knowledge in networks as exploitation and exploration. On the other hand, Apache Hadoop is a digital service platform. Competencies in information technology (IT) are an essential part of IT capabilities, as they facilities intra- and inter-organisational business processes as well knowledge management through the implementation of IT (Zhu, 2004). To build on the resources of information technology (IT) firms have to "explore and exploit" their IT capabilities. The "explore and exploit" process is compatible with Levinthal and March's (1993) concept about how firms pursue either productivity extraction or invention discovery, or both concurrently (Ortiz de Guinea and Raymond, 2019). It is the case of Apache Hadoop creators and firms using the Apache Hadoop to solve internal problems.

3.4. STRUCTURAL AND CONTEXTUAL AMBIDETERITY

The introduction of the chapter highlighted the concept of ambidexterity. Organisational ambidexterity is widely discussed in the management literature (Gibson and Birkinshaw, 2004; Gupta et al., 2006; Lavie and Rosenkopf, 2006; O'Reilly and Tushman, 2008; Raisch et al., 2009) (in Tushman et al., 2010). Organisational ambidexterity is described as the ability of firms to explore new service and products with new skills and exploit current service and products with existing knowledge (Andriopoulos & Lewis, 2009). The organisational literature discusses ambidexterity two folds: structural and contextual ambidexterity (Gibson and Birkinshaw, 2004).

Structural ambidexterity divides the explorative from the exploitative business unit in firms. The units are joined together by a senior manager. They have different unit's managers, processes, structures, and cultures (O'Reilly & Tushman, 2004). The senior managers are responsible for unifying the two separate units of the firms (Benner & Tushman, 2003). Contextual ambidexterity seeks to explore and exploit service and products simultaneous within a single business unit. The design stimulates the division of time spent on each task by individuals. Contextual ambidexterity enables employees to dynamically and flexibly determine the value of their activities for service and product development (Gibson & Birkinshaw, 2004). Structural ambidexterity is associated with short-term gains, contextual with long-term benefits. Structural ambidexterity handles the contradictory between explorative and exploitative in service and product development in separate business units. Contextual ambidexterity manages the disparities with the ability of multi-tasking by the employees (Gibson & Birkinshaw, 2004). Table 25 summarises the differences between structural and contextual.

Table 28. Structural Ambidexterity and Contextual Ambidexterity in Firms

Characteristics	Structural	Contextual
Units' Alignment	The exploitation and exploration of new service and product are divided in business units	Individuals divide their time between the exploitation and exploration of new service and products.
Decision Making	At the top of organisations	On the front line (at business units level)
The role of Management	The management is responsible to structure the firm to answer to the trade-offs between alignment and adaptability	Individual employees can decide where their contributions create more value in the context they act
Employees Responsibility	Clearly Defined	Flexible
Skills	Specialised	Broader

Source: Gibson & Birkinshaw, 2004.

3.5. INNOVATION STREAMS

Similarly, to ambidexterity, there is a establish frame of research focusing on the management of both efficiency and flexibility in organisational studies. Thompson (1967) discussed the paradox of administrations in balancing efficiency and flexibility. Abernathy (1978), in his work on the auto-industry, discussed that sustained performance was built upon the capability of organisations to move down learning curves as well as create new ones. Weick (1979) considered the topic as 'hypocritical organisations'. Quinn and Cameron (1988) introduced the notion of paradox. The authors suggested that organisations need to operate in multiple time frames and learning modes (Tushman et al., 2010). A more contemporary work by March (1991) claims that sustained performance is linked to the ability of firms to balance out exploitation and exploration processes. March (1991) has inspired other authors to research the subject (Eisenhardt and Martin, 2000; Rivkin and Siggelkow, 2003; He and Wong, 2004; Gilbert, 2005; O'Reilly and Tushman, 2008; Andriopoulous and Lewis, 2009).

One way to mobilise 'exploitation and exploration' in products and service development is the concept of innovation streams (Tushman and Smith, 2002; Benner and Tushman, 2003; O'Reilly and Tushman, 2008; Tushman et al., 2010). Innovation streams describe patterns of innovation. Innovation streams build on and extend products and service. But innovation streams may as well destroy the very products and service that was responsible for the firm's technological and economic achievements. Innovation streams shift the focus away from isolated innovations to patterns of innovation that evolve. Innovation streams advocates that success is built upon the firm's ambidexterity. Meaning, the ability of firms to learning anew as well as incrementally evolve. Successful firms have the skill to grow through steady exploitation of products and service development as well as through exploring across different types of innovation (Tushman et al., 2020). Innovation streams consist of continuous incremental innovation in existing products and service in conjunction with at least one non-incremental innovation (Smith and Tushman, 2005).

Scholars (He and Wong, 2004; O'Reilly and Tushman, 2008; Raisch and Birkinshaw, 2008) have highlighted an organisation's capacity to engage in exploratory and exploitative activities at the same time. Empirical data strongly suggest that organisational ambidexterity may enhance short-term performance and long-term survival rates in dynamically competitive situations (Junni et al., 2013). This is because ambidextrous businesses are more adept at reacting to disruptive new business models and emerging technology than other types of businesses (Birkinshaw et al., 2016; Hill and Birkinshaw, 2014). Ambidexterity in organisations does not result from the concurrent pursuit of exploratory and exploitative activities but rather from the effective mobilisation and integration of such actions or behaviours to generate novel combinations within organisations (Jansen et al., 2009; O'Reilly and Tushman, 2008). O'Reilly and Tushman (2013) pointed out that organisational ambidexterity demands collective interplay that develops and integrates previously disconnected ideas and knowledge in new directions. As a result, ambidexterity is expected to be implemented at several organisational levels (Kassotaki et al., 2019), motivating researchers to emphasise the importance of multi-level analysis (Jansen et al., 2012; Raisch et al., 2009) (in Tarba et al., 2020).

The combination of continuous learning with building on the past simultaneously bring about technological alternatives from which senior managers make strategic bets (Duncan, 1976; Tushman and O'Reilly, 1997). Innovation streams use the competences of firms to manage both incremental as well as discontinuous innovation. Organisational inertia must be combined with discontinuous organisational changes (Sastry, 1997; Romanelli and

Tushman, 1994; Gavetti and Levinthal, 2000). The notion of innovation streams can be used as a tool to illustrate how technology may evolve over time (Tushman & O'Reilly, 1997; Tushman & Smith, 2002).

According to the literature on strategy, effective organisations in dynamic settings are ambidextrous: They are efficient in meeting today's needs while being adaptable to future changes in the environment (Derbyshire, 2014; Enkel et al., 2017; Tushman and O'Reilly, 1996). Although beneficial, ambidextrous organisations often experience internal organisational conflicts between two strategies, such as efficiency and flexibility (Adler et al., 1999), or alignment and adaptation (Andriopoulos and Lewis, 2009; Gibson and Birkinshaw, 2004) (in Lin and Ho, 2021).

3.5.1. Technological Cycles

Abernathy (1978) and Abernathy and Utterback (1978) were the pioneers in linking dominant design and organisational evolution. Dominant designs and its ramifications on industry and organisational evolution have influenced several research streams since the late seventies. Dominant design strikes as core in both technological as well as organisational evolution. Scholars of different disciplines have discussed the impact of technology in forming organisational outcomes.

Innovations fuel technological progress . Disruptive innovations, which have a major and far-reaching influence on markets and society, are among the most important sorts (Danneels 2006; Yu and Hang 2010). Disruptive innovations have a different impact on markets depending on how they evolve compared to previous inventions. In this context, ideas based on processes of competitive substitution of new technology for an old one may be used to explain technological progress (Fisher and Pry, 1971; Sahal, 1981; Utterback, Pistorius, and Yilmaz, 2020). According to Berg, Wustmans, and Bröring (2019), the dynamic rivalry between technologies may lead to a new technology's domination over old technologies (Lin and Ho, 2021).

Some scholar (Mensch, 1979; Sahal, 1981; Dutton and Thomas, 1985) have discussed technological evolution and change in many different ways. Gilfillan (1935) explained technological change is an attribute to historical necessity in his studies on sail for ships. Schmookler (1966) and Merton (1969) associate technological progress to economic demand and growth. But there are strong indications from different technologies that none of the views on its own illustrates the complexity of technological change. Instead, technology appears to unfold in response to the interplay of history, individuals, and market demand. According to Morison (1966) and Sahal (1981), technological change is an exercise of variety and chance, along with structure and patterns. Technological change has been studied in case studies across a variety of industries and technological progress establish an evolutionary system underlined by discontinuous change. Important technological discoveries are seldom and are motivated by individual genius. They result in a span of technological ferment. As a new product or service comes to fruition (substituting previous ones) alternative products and service compete for market dominance. Trial and error and competition endure within a product or service class until a dominant design materialises as a fusion of several proven solutions (Utterback and Abernathy, 1975; Abernathy, 1978) (in Tushman and Anderson, 1986).

Scholars (Henderson and Clark, 1990 and Barley, 1986) share the observation that technology is one of the significant drivers of industrial, strategic and organisational change (in Tushman and Anderson, 1986). Single dominant designs (Abernathy and Clark, 1985; Tushman and Anderson, 1986; Teece, 1986; Suarez and Utterback 1995; Kodama, 1995;

Christensen and Bower, 1996) have dictated the narrative on the subject to the end of 1980s. In sharp contrast to single dominant designs, Tushman and Murmann (1998) introduced the notion of technology cycles. Technological cycles are spans of variation (eras of ferment) presented by technological discontinuities that finish with the draft of a dominant design. Dominant designs precede in the course of incremental change that is, in turn, interrupted by ensuing technological discontinuities, and the next cycle of change, selection and retention (Anderson and Tushman, 1990).

Dominant designs can be seen as the indicator for a subsequent era of ferment and incremental change. Once a dominant design materialises the ambivalence between designs disappears, and ensuing technical progress expands the selected variant. Anderson and Tushman (1990) have built on the literature on the history of technology. The authors do not observe a dominant design as the batch of subsystems that institute dominance. The authors look at dominant designs from the perspective of how they develop out of an evolutionary process defined by variations and selection. The process eventually leads to a retention period. From the standpoint of evolution, dominant designs are not compelled by technical or economic excel. Dominant designs are galvanised by sociopolitical/institutional means of negotiation and the act of finding a middle ground that accommodates the different interest of the parties involved. The negotiation is driven by economic and technical constraints (in Rosenkopf and Tushman, 1998). Whereas social, political and institutional forces create, dominant designs technology drives ensuing technological evolution. The cycles restart at the next technological discontinuity (Tushman and Murmann, 98).

Anderson and Tushman (1990) build the concept of 'cyclical model of technological change' on three frames of literature – the comprehensive review of technological evolution by Basalla (1988) and the work on SCOT by Pinch and Bijker (1987). The works by David

(1985/1987), Hughes (1983/1987), Abernathy (1978) and Jenkins (1975). Finally, the publications on the sociology of technology (Constant, 1980/1987; Landau, 1984; Tushman and Anderson, 1986, Barley, 1986).

3.5.2. Dominant Design

Dominant designs and ensuing technological discontinuities characterise technology cycles (Anderson and Tushman, 1990; Van de Ven and Garud, 1993). Abernathy and Utterback (1978) have explored the relationship between product and process innovation and looked at dominant designs as the condition driving industries to move from custom-made to a standardised-product manufacturing system. Dominant designs meet the needs of wide-ranging users. Most importantly, it is not a radical innovation but rather a creative combination of prior innovations. In line with the concept of the evolutionary approach, Anderson and Tushman (1990) remarked that dominant designs could only be recognised in the aftermath and a dominant design is established when a single design dominates over 50 % of sales (Tushman and Murmann, 1998). Table 29 illustrates the literature on dominant design based on Tushman and Murmann (1998) and Table 30 extends Tushman and Murmann (1998) with categories and factors for design dominance based upon van de Kaa and Kamp (2021).

Technology evolves along specific paths dependent on the technological paradigm, a criterion for assessing various solutions (Dosi, 1982), characterised as a framework of recognition that technology professionals share (Constant, 1980). However, identifying a worldwide optimum solution to a technical challenge is difficult due to multiple restrictions such as geographical and industry features. Consequently, technological progress is a cumulative process influenced by technological paradigms and arises from specialised paths. According to Dosi and Nelson (2010), if the technological paradigm shifts as a result of the introduction of a new technological source, the technological trajectory shifts as well, and thus the technological development path will exhibit complicated patterns as technological progress and the technological trajectory change, and changes in the technological paradigm overlap as a result of the emergence of the new source (Kim et al., 2021).

The dominant design is a collection of primary components and core ideas that do not differ considerably from one product architecture to the next and enable the demands of the mainstream market to be addressed (Abernathy and Utterback, 1978). A design like this dramatically decreases the number of performance criteria that a new product must meet by including many of them within the design itself (Utterback, 1994). Rather than increasing technical performance or a single attribute, the dominant design incorporates a set of functional traits that meet the needs of the mainstream market (Teece, 1986). According to Utterback and Suarez (1993), it is the consequence of a technological trajectory that involves a sequence of technical decisions concerning the product that are constrained by prior technical decisions and the growth of consumer preferences (Fernandez and Valle, 2018).

Campbell (1969) argues that for a variation and selection to consolidate in an evolutionary process, a retention mechanism is needed. The author believes that a successful variation needs to be perpetuated and reproduced. A dominant design is the second milestone in a technology cycle. It closes an era of ferment and is a single architecture that ratifies its dominance in a product or service class (Abernathy, 1978; Sahal, 1981). Table 29 summarises the discussion on dominant designs in the innovation literature based on Tushman and Murmann (1998) and Table 30 extends it with categories and factors for dominant design based on van de Kaa and Kamp (2021).

Table 29. Discussion on Dominant Designs in the Innovation Literature

Discipline	Discussion	Literature
Economy	shifting industry structures	Abernathy and Clark, 1985; Klepper, 1996; Langlois and Robertson, 1992
Strategy	product class and firm performance	Henderson, 1995; Teece and Pisano, 1994; McGrath et al., 1992; Prahalad and Hamel, 1994
Organisational	entry/exit rates and organizational fate	Wade, 1996; Podolny and Stuart, 1996; Baum et al., 1995; Rosenkopf and Tushman
Technology Management	to shifts in innovation types, firm performance, and industry structure	Iansiti and Clark, 1995; Utterback, 1994; Suarez and Utterback, 1995

Source: Tushman and Murmann, 1998.

Once a dominant design crystallises further technological advancements constitute of incremental improvements. The enhancements expand the standard and the technical system; the system grows into something more stable as one design expresses its dominance. The concept of dominant design has been used by several scholars to explain technological evolution. For instance, Utterback and Abernathy (1975) argued dominant designs are one of the fundamentals in the evolution of an industry and the transition from the state of instability to a state of stability. Clark (1985) and Henderson and Clark (1990) embraced and drawn-out the concept. Dosi (1984) and Nelson and Winter (1982) believed trajectories of paradigms form technological progress. It can be said that the work of Kuhn somewhat influenced the authors. Kuhn (1962) tried to explain scientific revolution as a trajectory defined by a paradigm – exceptional innovations put an end to the paradigm. Finally, Sahal (1981) claimed that certain designs what he would call ‘technological guideposts’, establish a model for ensuring technological progress. The technological guidepost once is in a firm position; innovation follows by incremental modification of the basic design (Anderson and Tushman, 1990).

Table 30. Categories and Factors for Dominant Design

Category	Explanation	Factors
Characteristics of the format supporter	The complementary assets and resources that technology supporters require to participate in a design dominance competition	Financial stability Credibility and brand reputation dominance in terms of operations, Orientation to learning
Characteristics of the format	Aspects of the design that help it outperform others in terms of technology.	Technological superiority, Compatibility, Complementary goods, Flexibility
Format support strategy	The technological supporters' strategic manoeuvring required to compete for design supremacy	Appropriateness, entrance time, marketing, Preemption of valuables, Commitment
Other stakeholders	Other stakeholders may impact the result of the design dominance competition.	Previous installed base, Big Fish, Regulator, Network of stakeholders
Market characteristics	Market characteristics that influence a technology's possibilities of market domination.	Bandwagon effect, Network externalities, Rate of change, Switching costs

Source: van de Kaa and Kamp, 2021.

The dominant design dictates its platform and drives technical progress. Following the introduction of a dominating design, the technological trajectory may alter dramatically.

Dominant design is characterised as the tipping point for opening industrial innovations via standardisation, allowing manufacturers to establish mass-manufacturing systems (Abernathy and Utterback, 1978; Murmann and Frenken, 2006). The dominant design on product development is validated by comparing technological trajectories before and after its introduction (Kim et al., 2021).

3.6. INNOVATION CLASSES

The research on 'technologies' follows similar characteristics to other researches in 'science' – they can also be discussed as 'technological paradigms' as they function the same way as "scientific paradigms". Technological paradigms try to depict continuous changes and discontinuous technological innovation. Continuous change is often associated with advancements along a technical path symbolised by a technological paradigm. At the same time, discontinuities are linked to the emergence of a new paradigm. The ties between economic growth and change and technical progress are apparent and well-studied in economics and well accepted. There is significant endeavour in the economic literature to contextualise the essences of inventive activities and the common elements among a variety of inventions and/or innovations[3]. Theory of innovation should be able to explain not only "incremental" technical advancements on an existing product/process, but, first and foremost, to interpret major and minor technological changes (Dosi, 1982).

Radical and incremental specify different classes of technological process innovations. Radical innovation is underlying changes that characterise revolutionary changes in technology. Radical innovation symbolises an apparent distancing from current practices (Duchesneau, Cohn and Dutton, 1979; Ettlie 1983). They are opposed to incremental innovations – small enhancements or trivial alterations in current technology (Munson and

3 Innovation is associated with the commercial introduction of a new product or a new combination. Invention is linked to the realm of science and technology (in Perez 2010).

Pelz, 1979). The notable contrast between radical and incremental is the level of originality in the technological process embodied in the innovation and hence, the level of new knowledge embedded in the innovation. The difference fits in and is in line with researches that characterise technology as a knowledge component (Dutton and Thomas 1985). The difference between radical and incremental should be viewed as a continuum – a path from incremental to radical innovations (Hage 1980). According to McGrath (2001), dominators of incremental innovations are refinement, production, efficiency and execution. In contrast, the ones for radical innovations are search, variation, experimentation, flexibility and risk-taking (Brettel et al., 2012).

3.6.1. Incremental Innovation

Incremental innovation improves and widens products and service price performance and/or product design by continuously exploiting a technological path (Dosi, 1982; Rosenkopf and Nerkar, 2001). The innovations are usually an addition to the current products/service portfolios of firms and follow a logical technical progression by extending processes and/or designs of products and/or service. It suggests small changes, exploits already confirmed designs and usually strengthens the authority of organisations (Nelson and Winter, 1982; Ettlie et al., 1984; Dewar and Dutton, 1986; Tushman and Anderson, 1986).

These changes can have a positive impact regarding competitiveness, as they enhance functionalities and/or lower costs and improve efficiency (Sen and Ghandforoush, 2011). The type of innovation upgrades the current designs of components – and should be seen as minor changes rather than radical improvements. The modifications are based upon an organisation's know-how in component technology and/or service design within a proven architecture (Christensen, 1997), associated with trivial changes to the current design (Tushman & Romanelli, 1985). It is often understood as a gradual development (Christensen, 1997). Continuous innovation is vital in software because it is a theoretical illustration of the constant updates of a programme. Many upgrades are available almost daily. Some of them may address some bugs and improve some lines of code – making the software more adaptable for future changes. When some software is updated from 1.7 to 2.0, the interface might be different, or the location of a very particular tool may change. The knowledge can be built upon already existing knowledge. Innovation based economies involve ethos and considerable financial investment. The earnings are often long term (Talke 2007). It is understood that innovation is one of the main drives for economic growth, and incremental plays a crucial role in the diffusion and marketability of products and service (Sen and Ghandforoush, 2011).

Practitioners and researchers acknowledge incremental innovation as the core of organisational innovation success (Oerlemans et al., 2013). Incremental innovation is concerned with the continuous enhancement of existing technologies in terms of design, function, features, and so on to suit the demands of current consumers (Margaret and Nathaniel, 2019). At the organisational level, incremental innovations can be used to achieve a competitive edge and act as shock absorbers, allowing organisations to make modifications in reaction to changes in their surroundings (Un, 2010). According to Robertson et al. (2012), incremental adjustments will result in long-term competitive advantages for inventive organisations. In contrast, firms that underestimate the relevance of incremental gains over time may suffer fatal or near-fatal setbacks (Shi et al., 2020).

Incremental innovations unfold by widening current knowledge and expertise and outdo established designs. Since incremental knowledge consists of existent knowledge and emphasises current skills and processes by elevating the capabilities, putting into use current know-how and concentrating on current activities in existing domains – they are also called exploitative innovations (Lewin et al., 1999). Organisations that embrace a responsive market orientation exploits the needs of its customers for the development of new products (Grinstein, 2008). Exploitative mode of operation correlates with organisational behaviour – usually sequential and incremental. The ways are centred on problems or circumstances that are within the conventional capacities of the firms' actions (Slater and Narver 1995; Baker and Sinkula 2007). According to Li et al. (2008), firms responding to market inclinations may profit from continually improving or upgrading the original product and/or service for their customers through incremental innovation (Brettel, 2012).

Cumulative expertise is singular. When it outperforms the competition, it may drive a competitive advantage. But, cumulative expertise results in rigidities' (Leonard-Barton, 1995) or fall into 'competency traps' (Levitt and March, 1988). Incremental improvement of a product and service category becomes increasingly specific - less appropriate to other classes. It becomes even more apparent when technology disrupts the incremental path. Consequently, the more an organisation concentrate on one particular technology, the more significant the impact on the subsequent technological evolution is on the organisation (Rosenkopf and Nekar, 2001).

3.6.2. Architectural Innovation

A new technology that stimulates new ties to markets and customers are typically exemplified by the formation of new trends and/or the remodelling of old ones. Such innovations shape the main structure of products and processes and determine the technical paths that will lead to future advancements. They dictate the architecture of an industry and its agenda within which rivalry will take place and mature. They put forward new industries and/or reshape old ones. Architectural innovation comprises processes and organisational changes that are very different from the established ones (Abernathy and Clark, 1985).

Based on the frame of literature, Henderson and Clark (1990) have coined the concept of architectural innovation as a source to create a competitive advantage. Architectural innovations were defined “as those that exchange how the components are linked together while leaving the core design concepts (and thus the basic knowledge underlying the components) as untouched” (in Henderson and Clark, 1990, p. 10). And, a component described as “a physically distinct portion of the product that embodies a core design concept ... and performs a well-defined function” (in Henderson and Clark, 1990, p.11.). The authors suggested a structure which divides product and service innovation into four incremental, modular, architectural and radical. The four types of innovation portray the effect of innovation on current technologies underlying the components or architecture of products (Bozgodan et al., 1998).

Innovation, whether initiated by a rival or pushed by a supplier (Christensen, 1997; Wagner and Bode, 2014), can be disruptive to a firm's existing operations and market position (Christensen and Raynor, 2003). To respond to exogenous innovation, organisations often need to adjust their operations (Craighead et al., 2009; Zimmermann et al., 2016), which necessitates changes to their supply chains and interactions with critical external customers and suppliers (Chae et al., 2020; Peng et al., 2013). More fundamentally, a significant external innovation might drive enterprises to reconsider and ultimately shift their make-buy sourcing strategies (Argyres et al., 2019; Bigelow, 2019). As an industry matures, a dominant design style often emerges (Abernathy and Utterback, 1978; Anderson and Tuchman, 1990). According to Gallagher (2007) and Murmann and Frenken (2006), organisations then optimise various combinations of critical product characteristics and components until the most desired or optimum product solution emerges as a modular standardisation (Park and Tangpong, 2019).

Architectural innovation consists of reorganising known parts – components – into new patterns (architectures) to improve performance in one or both parameters. According to Henderson and Clark (1990), "architectural innovation destroys the usefulness of [the non-innovator's] architectural knowledge, but preserves the usefulness of its knowledge about the product's components (p. 10)." Architectural innovation relies upon the innovator's superior architectural knowledge. Ulrich (1995, p. 419) characterises a product architecture as "the scheme by which the function of a product is allocated to physical components." Architectural knowledge deals with the operations of the system and how the components of the system support the operations. In a like manner Baldwin and Clark (2000, p. 77) describe a system's architecture as the "modules that will be part of the system, and what their roles will be." (in Baldwin and Clark, 2006).

Innovations that modify the ways in which the components of a product are connected together, leaving their main designs unaffected, are architectural innovations. They terminate a business's architectural savoir-faire but still protect the important understanding about a product's smallest units. These innovations emphasise the differences between a product/service as a system and a product/service as a set of components, suggesting that product/service design involves two types of expertise. It involves knowledge of all components, the main design choices, and how these small units are linked together and implemented. It also calls for architectural knowledge, or knowledge about the means by which the small parts are combined together as a whole. It is the rearrangement of an accepted system to connect together components in new manners. The major argument is that the main design behind each component – and the related technical and business understanding – persists. Architectural innovations take advantage of many established core designs and concepts in a new "architecture" and stress the relationship between components (Henderson and Clark, 1990).

The linkages between the product's components become more modular and standardised (Baldwin and Clark, 2000). Firms no longer need custom-designed components, which diminishes the necessity of transaction cost management (Christensen et al., 2002). The competitive landscape now changes toward overall cost reduction, focusing on efficiency (Utterback, 1994). Additionally, as the product becomes more modular and standardised throughout the industry, specialist providers can easily infiltrate the market. This line of study has advanced our knowledge of how organisations might strategically shift their source preferences over the innovation time horizon (Park and Tangpong, 2019).

As discussed in this section, a stream of literature with its roots in the 1980s and 1990 underscored the significance of the firm’s architecture in deciding the winners and losers in the commercialisation of emerging technologies. The literature describes the elements and how these elements communicate with each other. In recent years, theorists and practitioners have become more familiar with the idea that industries, as structures, may also have architectures. Industrial architectures define the existence and degree of specialisation of the participants in the market (or organisational boundaries) and the arrangements of the relationship between them. Academic studies over the last decades have demonstrated that there is a close correlation between the architecture of the market and the architecture of physical goods and technologies. The computer industry has developed from vertical to horizontal architectures due to the modular technical architectures of computers. Industrial architectures characterised by specialised or networked organisations embodies product/technology architectures that integrate well-defined interfaces between components technologies or high degrees of modularity (Pisano and Teece, 2007).

Architectural innovation incorporates the disruption of technology and business structures. Digital photography can be used here as an excellent example. Firms such as

Kodak and Polaroid, entering the modern world required learning entirely new expertise in state-of-the-art computing, camera architecture, software and display technology. It also represented finding new ways to prosper from cameras rather than from disposables (film, paper, processing chemicals and services). The innovation plan of firms should define how the various types of technologies work into the firm's strategy and the resources to be devoted to each. Radical, disruptive, and architectural innovation are perceived as the to growth, and incremental innovation is denigrated as myopic at best and suicidal at worst (Pisano, 2015).

3.6.3. Discontinous Innonation

Discontinuous innovation is crucial for competitive advantage and may lead to a firm's growth and profitability (Bohlen, 1964; Calantone et al., 1988; Kleinschmidt, 1991; Everett, 1962). There is a large frame of literature on discontinuous innovation (Green et al., 1995; Hage, 1980; Lee and Na, 1985; Meyers et al., 1989; Nyström, 1985). The expression of "discontinuous innovation" often refers to radically new products and service (Meyers et al., 1989). Product and service innovation are discussed in three dimensions: product benefits, technological capabilities, and consumption or usage patterns (in Veryzer, 1998). Table 27 summarises the aspect of product innovation based on Veryzer (1998).

Discontinuous innovations demand flexible and responsive approaches from firms as future developments are challenging to forecast (McKelvey, 2004). Discontinuous innovations are new technology, products, or business models that significantly deviate from an industry's existing state of the art (Birkinshaw, 2007). Such innovations often occur due to the introduction of new technology to a market. Usually, the technology is already used in a related sector but is customised to meet the new environment (Bergek et al., 2013; Magnusson et al., 2003). According to Picaud-Bello et al. (2019), discontinuous innovations need organisations to establish and maintain inter-organisational partnerships with various external stakeholders, including suppliers. According to some authors (Bessant et al., 2005; Lynn et al., 1996; Phillips et al., 2006b), current studies in innovation management demonstrate that discontinuous innovation emphasises the development of new capabilities both within the firm and with other firms, such as "non-linear, highly exploratory, and experimental organisational processes" (Picaud-Bello et al., 2019).

Innovations are usually discussed as the development of a product, service, or process. They can be understood from an evolutionary perspective or "continuous", to revolutionary or "discontinuous". Expressions such as "radical", "breakthrough", "revolutionary", "really new", "game-changing" and "boundary-expanding" have all been applied to describe products/services that comprise considerable deviations from current products'/service's logical progression. Discontinuous innovations symbolise radical new products/service. These new products/services consist of radical changes in terms of consumer understanding and usage (Myers and Tucker, 1989). Discontinuous innovations not only encompass unique progress but also commercialisation difficulties due to the extreme degree of ambiguity regarding their technical and market viability (Veryzer, 1998).

Appropriating Thomas Kuhn's (1962) idea of scientific paradigms, Dosi (1982) differentiates between "normal" types of technological development – which push an innovation along a clear, determined track – and the establishment of new "technological paradigms". Dosi describes a technological paradigm as the configuration of designated technological efforts, grounded upon determined beliefs of science and chosen technologies. New paradigms symbolise discontinuities in advancement articulated in earlier paradigms. Rice et al. (2002) Explain discontinuous innovations as "game changers". Kaplan (1999) argues that incremental innovations are important for defending incomes, but for major growth over a long period of time it takes discontinuous innovations. Tushman and O'Reilly (1997) suggest that discontinuous innovations encourage firms to abandon the past in order to develop new processes that can deliver a new edge to customers. Christensen (1997), Hamel and Prahalad (1994), and Utterback (1996) see discontinuous innovation as related to "disruptive technologies", "discontinuities" or "radical innovation". These innovations allow whole new industries and markets to arise, renovate or vanish. O'Connor (1998) explains discontinuous innovations as the development of new business channels – new for organisations and the marketplace these firms are in. Discontinuous innovations are comprised of a major technical variation in a product's main subsystem (Dosi, 1982; Tushman and Smith, 2002) that has fundamental impact on the product and/or service as whole (Tushman and Murmann, 1998).

Discontinuous innovation changes the liaisons between customers and suppliers, or even rearrange the marketplace economics. Discontinuous innovation replaces current products, and often shape completely product categories never seen before. Traditionally, discontinuous innovation is characterised by high technical and market uncertainty as opposed to incremental innovations – where technical and market uncertainty are relatively low. Government regulation and social trades can generate opportunities but new technologies are often the fundamental of discontinuous innovation. Discontinuous innovation is frequently depicted as technological leaps that assists organisations redraft rules or develop whole new industries. According to Ahuja and Lampert (2001) and Gatignon et al. (2002), discontinuous innovation consists of paramount technical changes in products and service. The types of innovation drive cascading effects throughout the product (Tushman and Murmann, 1998).

3.7. CONCLUSION

The chapter introduced the concept of innovation streams. Ambidexterity in combination with the three-dimensional way (incremental, architectural and discontinuous) to discuss innovation makes innovation streams presents a clear advantage to understanding a platform (Apache Hadoop) that consists of many different modules, architectures and markets. Apache Hadoop distributors are ambidextrous by design and the organisational literature discusses ambidexterity two folds: structural and contextual ambidexterity.

Similarly, to ambidexterity, there is a establish frame of literature in management science discussing both efficiency and flexibility in organisations. One way to mobilise 'efficiency and flexibility' as 'exploitation and exploration' in products and service development is the concept of innovation streams. Innovation streams use the competences of firms to manage both incremental as well as discontinuous innovation and can be used as a tool to illustrate how technology may evolve over time.

Dominant designs can be seen as the indicator for a subsequent era of ferment and incremental change. Once a dominant design materialises the ambivalence between designs disappears, and ensuing technical progress expands the selected variant. Dominant designs and its ramifications on industry and organisational evolution have influenced several research streams since the late seventies. Dominant design strikes as core in both technological as well as organisational evolution. Scholars of different disciplines have discussed the impact of technology in forming organisational outcomes.

Continuous change is often associated with advancements along a technical path symbolised by a technological paradigm. At the same time, discontinuities are linked to the emergence of a new paradigm. Incremental innovation improves and widens products and service price performance and/or product design by continuously exploiting a technological path. Architectural innovation consists of reorganising known parts – components – into new patterns (architectures) to improve performance in one or both parameters. Discontinuous innovation is crucial for competitive advantage and may lead to a firm's growth and profitability. Discontinuous innovation changes the liaisons between customers and suppliers, or even rearrange the marketplace economics. Discontinuous innovation replaces current products, and often shape completely product categories never seen before.

Discontinuous and incremental innovation are extreme points along both dimensions. Discontinuous innovation produces a new dominant design and thus a new collection of fundamental designs embodied in components that are joined together in a new architecture. Incremental innovation refines and improves the design that has been developed. Improvements arise in individual elements, but the basic key architecture principles and the linkages between them remain the same.

CHAPTER 4. METHODOLOGY

4.1. INTRODUCTION

The thesis has so far specified the research scope and the theory through which the thesis is developed. This section illustrates the research design that is pursued to answer the two research questions:

- How do Apache Hadoop creators and consumers generate new innovation streams?

- How can innovation streams evolve from community base to enterprise platforms?

Research methodologies are characterised in various ways. The one followed in this work holds the belief that methodology is rationally engraved in theoretical perspectives which are in line with an epistemology of reference: theories on means of knowledge, and principles to validate them. A coherent way to inform social research is through the four-step method developed by Crotty (1998): (1) epistemology; (2) theoretical perspective, (3) methodology, (4) ad hoc research methods. Figure 3 illustrates the four steps formulating the rationale that advances from one stage to the other. Crotty's (1998) model is applied to improve the rigour of the research process in all fields of social science. The work also follows an additional methodological instrument for data collection, which seeks to ensure the quality and coherence of the research project. It is comprised of McCracken's (1998) "Long Interview Techniques" principles.

The method used in this work is narrative inquiry. Case histories are used to understand how innovation in the Apache Hadoop ecosystems takes place and how the ecosystem evolves. The primary tool used in the investigation was interviewing, and the data was analysed based on theoretical sampling. Semi-structured interviews were the primary vehicle to direct and script the narratives. The semi-structured interviews were critical in helping Apache Hadoop professionals to think about and share their practices as they happened while working on various Apache Hadoop projects. The histories told by the participants were fundamental to structure the investigation and to collect personal accounts of the Apache Hadoop professionals in Brazil.

Figure 3. *The Four Steps Method*

Source: Crotty, 1998.

This chapter is structured as follows. Section 4.2. Epistemology discusses philosophical positioning. 4.3. Qualitative Research deals with the theoretical perspective and involves the use of qualitative research in ISstudies. 4.4. Narrative Inquiry introduces the adoption of case studies in the information systems. 4.5. and 4.6. describe the Data Collection and Data Analysis. 4.7. offers some limitations before the chapter is closed with 4.8. Conclusion.

4.2. EPISTEMOLOGY: SOCIAL CONSTRUCTIVISM

Constructivism, often referred to as interpretivism, is a way to understand the world. The meaning of objects is subjective and constructed around the experiences made about them. They can vary in character, and imply that the observer looks for complexities of the object rather than simplifying their meaning. Researchers base their views around the narratives of others, and the subjectivity is negotiated in a social and historical setting (Creswell, 2013). Defenders of constructivism aim at comprehending the world through the experiences made by others who lived them. The main goal of a constructivist is to define the essential condition of the understanding of others, including the reality lived by others, and the meaning given to specific objects in different environments (Schwandt, 1994). Constructivism intends to comprehend the various constructions that people have when they try to provide the sense to object (Appleton and King, 2002). The constructivist and interpretivist advocate that to understand the meaning of the world, it is necessary to interpret it. The inquirer must shed light on the process that gave meaning to objects to social actors. This language is embedded in the way these social actors express themselves. To interpret the language of these social actors, the inquirer needs to reconstruct their constructions of different (Schwandt, 1994).

Constructivist researchers advocate that people give meaning to “facts”, actions and something known through the senses through complex interactions (Schwandt, 1994). Based on Guba and Lincoln (1982) and Lincoln and Guba (1985), Appleton and King (1994) give five fundamental principles that underpin the paradigm of constructivism: (1) reality and its elements (2) causality (3) unique context arising from the lack of generalisation (4) the connection between the inquirer and the object of inquiry (5) the effect of values on the process of study. Constructivists understand that there are opposing views and are aware that objective knowledge and truth is a matter of perspective – knowledge and truth are generated not in the mind. Constructivists accentuate the pluralistic and plastic attributes in life: pluralistic due to the different symbols and languages; plastic due the way reality is drawn to form and/or apt to deliberate acts of social actors (Schwandt, 1994). In “constructivism”, knowledge does not unravel from discoveries of the truths of objects but through the constructions made by social actors about objects.

4.3. THEORETICAL PERSPECTIVE: QUALITATIVE RESEARCH IN INFORMATION SYSTEMS

IS research deal 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 by both (Avgerou, 2000). There is tension in regards to the essence of IS research (Lee, 2001; Baskerville and Myers, 2002; Avison and Fitzgerald, 2003). Some scholars (Orlikowski and Iacono, 2001; Benbasat and Zmud, 2003) claim that the IS field is in disarray as to what the essences of the field are, yet Walsham (2012) argues that the diversity of IS is of inestimable value. It is the multiplicity of theoretical methods that have proposed answers from the overly technical to more philosophical questions (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). Table 31 illustrates some advantages and disadvantages of the epistemological paradigms.

Table 31. The Advantages and Disadvantages of the Epistemological Paradigms

	Advantages	Disadvantage
Positivism	Economical collection of large amounts of data Clear theoretical focus for the research at the outset Greater opportunity for the researcher to retain control of the research process Easily comparable data	Inflexible Weak at understanding social process Often does not discover the meaning attached to social phenomena
Interpretivism	Facilitates understanding of how and why Enables the research to be alive to changes which occur Good at understanding social processes Allows for complexity and context factors	Data collection can be time consuming Data analysis is challenging and can be complex Researchers have to live with uncertainty that clear pattern may not emerge Generally perceived as less credible by 'non-researchers'

Positivist research is based upon prior fixed relationships within the phenomena which are studied with structured means; such types of research are suitable primarily for studies trying to test theories. In contrast, interpretive studies purposely embrace a non-deterministic viewpoint and attempt to explore the phenomena of interest in its natural environment, consciously not enforcing any prior understanding of it (Orlikowski and Baroudi, 1991).

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(IS) have come under the spotlight in recent decades (Walsham, 1995). Qualitative research uses qualitative data, such as interviews, documents and participant observation, to understand and explain social phenomena (Benbasat et al., 1987; Eisenhardt, 1989). Bearing in mind the area of concern, the research questions 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).

4.4. METHODOLOGY: NARRATIVE INQUIRY IN INFORMATION SYSTEMS

The interpretive narrative has stimulated interest from social scientists in recent years (Abbott, 1990/1992; Finnegan, 1992; Hinchman and Hinchman, 1997; Mishler, 1995; Riessman, 1993; Somers, 1994). Interpretive narrative can be thought of as a way of organising a series of events into a whole, with each event's importance determined by its relationship to the whole. The term narrative has been used by qualitative researchers in various ways. Narrative inquiry refers to a discourse type in which incidents and happenings are woven together in time by a story. Czarniawska (2004, p.17) defines it as a specific type of qualitative design in which "narrative is understood as a spoken or written text giving an account of events and actions or series of events actions, chronologically connected" (in Creswell, 2013).

The narrative approach is a frame of reference focusing on the inquiry process, an analysis tool and a way of thinking (Connelly and Clandinin, 1990). Thus, the narrative is both a phenomenon and a method. According to Moen (2006), three fundamental arguments are often found in the narrative studies literature. The first argument is that humans arrange their world encounters into narratives. Second, narrative scholars believe that the stories told are influenced by the individual's past and current views and beliefs, the individuals about whom the stories are told and where they are told. The third point, which is closely related to the second, is that narratives are multivoiced (Czarniawska and Joerges, 1995).

The work stands on the shoulders of other authors in IS research. For instance, Bartis and Mitev (2008) track how different stakeholders assigned different meanings to IS infrastructures. A narrative methodology was used to analyse the rhetorical strategies used. The authors employ relevant social groups and the social construction of technology to explore the many interests that influence organisational dynamics. Dalcher and Dreevin (2003) utilise narrative to understand IS development failures in many domains and countries. According to the authors, failures, like other organisational activities, are built on stories. The verbal medium is critical for understanding behaviour inside organisations and systems. Therefore researchers must gather histories about what happens, based on practice. One of the most difficult challenges for researchers of complex phenomena is determining what to do when several versions and narratives exist rather than a single, well-understood version shared by all participants. Understanding failures frequently involves the untangling of complex webs of actions and events and emergent interaction patterns. Heidelberger and Uecker (2009) argue that a constructivist research methodology uses a narrative that acknowledges the researcher's personal experience as a legitimate research object. The scholars suggest that that narrative gives IS researchers a new instrument to increase research relevance and discourse. Tan and Hunter (2003) utilise narrative in their study on the factors influencing the career paths of IS professionals. Their study reported a deeper understanding of career path influences by integrating human and organisational characteristics and building an opinion on the data on the perceptions of IS professionals and their interactions with their environment.

Other authors within organisation studies have found that an interpretive approach is suitable for IS research (Kanungo, 1993; Walsham, 1993; Myers, 1995, 1997). This work uses narrative analysis to clarify and provide meaning to what stakeholders say. To understand how businesses externalise their ideas and commercialise their innovations, the diverse perspectives of different stakeholders must be combined. Following this rationale, the different stakeholders are detailled in section 5.7.1. Key professionals.

4.4.1. Choice Of Narrative Inquiry Method

As the work interpreted what was said by the participants, the underlying method of the work is interpretative. A constructionist perspective grounds the narrative analysis method; the goal is not to discover 'facts' as reported by participants but to understand the contexts within which people construct their views, factual or not. Bruner (1986) believes that human beings create and arrange reality in two fundamental ways: paradigms and narratives. On the one hand, evidence is sought through empirical verification in the paradigmatic method. Its objective is to make reality more transparent. According to Bruner (p. 12), "its language is regulated by requirements of consistency and noncontradiction". On the other hand, the narrative method emphasises telling stories that are culturally and chronologically connected. This method produces in-depth understandings as the participants can narrate their own stories – this is reflected in the case histories section . These personal experiences of specific topics are documented and analysed. Bruner (1990) further claims that the narrative approach entails documenting contextually rich and chronologically constrained stories. The phrase "contextually rich" refers to first-hand descriptions of vividly recalled experiences. When a story is structured with a beginning and an end, it is temporally bounded. While narrative inquiry can be used as a qualitative method, it is the researcher's responsibility to choose a strategy that enables this approach while providing structure to the data collection process. The researcher must select a design that allows this approach while providing a system to the data collection process (Tan and Hunter, 2003).

Interpretive analysis requires understanding how the research subjects relate to events and how the subjects describe the events. It, in turn, calls for 'imaginative reconstruction' or 'empathy' from the researchers. Narrative analysis is a subjective effort. A satisfactory interpretation of qualitative evidence, which enhances the understanding, needs some heuristic mechanisms that allow them to move beyond the apparent or most evident substance (Elliott, 2005). Narrative interpretation enables insight into two realms: (a) the realm of experience, in which speakers describe how they, as individuals, see particular occurrences and assign subjective meaning to these perceptions; and (b) the realm of storytelling strategies (or technologies) used to create sense (Bamberg, 2012). Czarniawska-Joerges (1995) argues that narratives are applied in three ways in organisational studies:

"In tales from the field" (in organisational studies written in a story-like construct).

"In tales of the field" (in organisational studies gathering corporate stories).

"In interpretive approaches" (organisational studies conceptualising administrative activities as story making and organisational theory as story reading).

In addition to Czarniawska-Joerges' (1995) "In interpretive approaches", the work follows other works in IS that have used interpretive narrative techniques. For instance, Boland and Day (1989) run a series of interviews to capture the experiences of IS designers. The authors argued that they had gained a better grasp of the IS designer's meaning during the design process. Davidson (1997) employed narrative methods to investigate social cognitive processes in IS development. By creating and maintaining a narrative account of the project's history, the participants could narrate their experiences, promoting a better understanding of the process (Tan and Hunter, 2003). The work also follows Chatman's (1978) "necessary components" of narrative. A narrative consists of the story (a chain of events), characters and discourse for the author. The discourse can be understood as the way these components are informed. In Hatman's own words (p. 19), '"Story is the "what" in the narrative that is depicted, discourse is the "how". Story is "what happens to whom" and discourse is "how the story is told"' (Webb and Mallon, 2007).

The topic considered in this study and the research questions through which it has been mobilised did not automatically influence a unique research method. However, a decisive logic behind the decision to apply interpretive narrative, which is associated with a somewhat lengthy attendance in the field, has been made clear by Rosen (1991). The author claims that understanding the process, the primary purpose of interpretivist research involves "getting inside the world" of those generating it. Thus, the interpretive narrative is driven by the need to penetrate the participants' world and understand the narratives that unfold through the participants' stories. In other words, the researcher must investigate people's points of view, which shape their social constructions, in the setting in which these are formed. According to social constructionism, the production of reality is continually moulded by a social dimension: to acquire a solid grip on these "facts", the researcher should comprehend well the social world to which they belong.

According to Rosen (1991), field studies are the best approach to interpretative research because they allow researchers to "get into the world" of people being researched. The generalisation of results, which, in positivism, is based on sampling and statistics, is structured along theoretical lines. Interpretivism does not generalise its findings based on a statistical sample. Instead, it applies its findings to broader, more widespread theoretical notions. Thus, interpretivism's generalisation progresses from empirical data to theory (Lee and Baskerville 2003; Yin 2003). Generalisation is discussed in more details in 4.5.2. Generalisation and Theoretical Saturation. Case histories are used as this study's strategy, with selected Apache Hadoop projects being explored. The aim is to illustrate the transition from exploitation to exploration and from discontinuity to a dominant design, and how these dynamic capabilities are critical for the evolutionary system of Apache Hadoop.

4.5. DATA COLLECTION

4.5.1. Interview Data Analysis Protocol

The research *followed Jacob and Furgerson's (2012) guidelines for conducting interviews. Table 32 summarises the ten-step process.*

Table 32. Guideline for conducting interviews (Jacob and Furgerson, 2012)

Interview Protocol Design	
The Use of Script	It enables you to give all of the critical information about your study and the essential rules about consent.
Informed Consent	Before starting the interview, give the participant a participant information page and an informed consent form to read and sign.
Basics	Begin by asking for some basic background information to help develop rapport and a comfortable setting for the participant.
Open-ended Questions	Open-ended questions give participants more opportunities to open up and elaborate on their experiences.
Question Informed by Research	Examine the literature on the issue before drafting your interview questions.
Gradual Increase of Complexity	One should not presume the participants share your knowledge. Abstraction might alienate your participant. On the other hand, asking more specific questions that participants can readily answer builds rapport and trust faster.
Use Prompts	Expect the respondent to elaborate on their own experiences and perspectives, but participants may require prompting. So plan out how you might assist someone to answer each of your open-ended questions (Jacob and Furgerson, 2012).
Revise Protocol	Depending on the participant's reaction and the interview's direction, you may need to re-arrange the questions.
Time Keeping	Remember that you ask others to share their experiences and time with you, so plan accordingly.
Pilot Test	By practising the questions aloud, you will get better acquainted with their sequence and flow while conducting the interviews for your data collection.

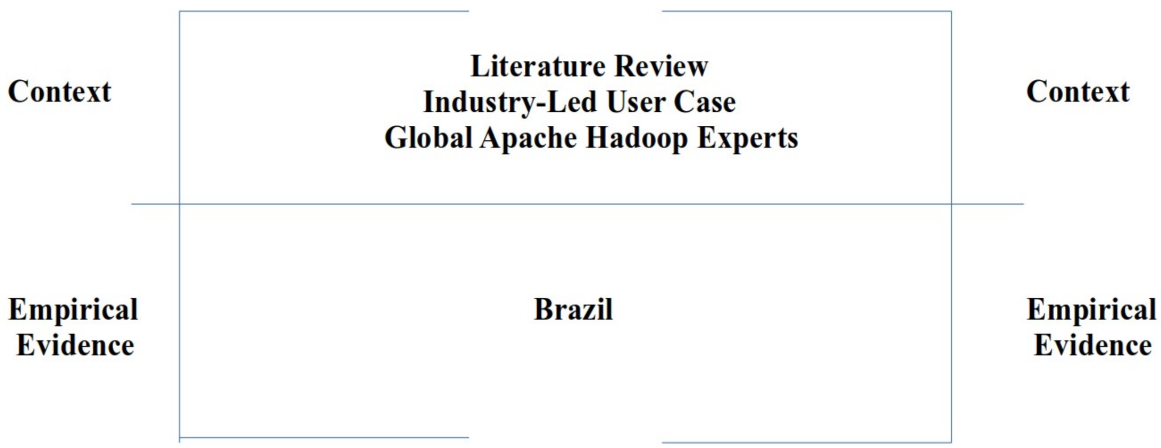
Source: Interview Design Protocol
(https://www.imperial.ac.uk/education-research/evaluation/tools-and-resources-for-evaluation/interviews/interview-protocol-design/)

4.5.2. Data Collection

The fieldwork took place in Brazil from October to December 2016. The Brazilian data comprise 25 semi-structured interviews. These participants are from 23 different organisations in four different cities. 14 firms were located in the great São Paulo area, a small sample of six were from three businesses in Rio de Janeiro, and the remaining participants were from Brasília (2) and Porto Alegre (1). The interviews lasted approximately one hour each, were all conducted via Skype, audio-recorded and later transcribed for analysis. All interviews were conducted in Portuguese, and the participants were found on LinkedIn.

In light of this study, a LinkedIn search for Brazilian firms working with the Apache Hadoop was conducted, with no restriction on industries. These firms innovate incrementally as well as architecturally and/or discontinuously, and have a LinkedIn group account. Most participants belong to groups in LinkedIn associated with the Apache Hadoop framework. The contact with the participants was a multiphase process: (1) an email was developed in Portuguese and a consent form in English (2) LinkedIn was scanned for suitable candidates. The requisites were: profession, years of experience and the types of organisation in which these professionals are employed; (3) if a person was found to be appropriate, they were invited to the LinkedIn network of the researcher. Once the invitation was accepted more personal details became available, including email address and mobile number.

Figure 4. *The Link Between Context and Detailed Analysis*



In cases where this information did not become available, the messages were sent through the LinkedIn platform. The initial response was very small, but this soon improved once the participants were contacted via mobile phone. In total, 150 participants were contacted: 45 agreed to be interviewed and 25 interviews were eventually conducted. The first phase gives insights that are already in the literature and were confirmed by the participants. The second phase discusses ideas that are already mentioned in the innovation literature, but to which I will add new observations. The third and final phase draws relevant conclusions that perhaps will need more systematic work. Finally, the link between the context and detailed analysis is illustrated in Figure 4.

Sampling in qualitative study design is often theory driven. Theoretical or purposive sampling involves the creation of interpretative assumptions from emerging data, and the selection of samples to test and expand on the theory. Theoretical or purposive sampling is seen in different formats in most qualitative inquiries involving interpretation. The relative balance depends on the topic of the study and the style of data analysis and interpretation chosen. One of the qualities of qualitative studies is its naturalistic essence: it examines real people in their natural environments rather than in artificial confinement. Sampling needs to consider the individual's qualities in addition to the temporal, spatial and situational forces; basically, the context of the study (Marshall, 1996).

Sampling in qualitative studies does not refer to achieving numerical significance or a particular sample size to establish what researchers interpret as significant (Merriam and Tisdell, 2016; Pope and Mays, 2006). Creating data in qualitative research means strategically appointing participants whose insights and experiences can contribute to, illuminate, and, potentially, help explain the phenomenon under study (Cleary et al., 2014). A key aspect of any qualitative study is sampling that helps researchers to identify similarities, discrepancies and anomalies in finding new and relevant data. Sampling in qualitative analysis is generally interpreted as focussing on the features of the cases or instances as important data sources for the investigation. A method that is widely used to collect data in qualitative research is purposive sampling, finding and strategically selecting information-rich cases pertaining to the topic of interest (Conlon et al., 2020). Accordingly, find a list of the participants and organisation in Appendix A.

The data were analysed based on theoretical sampling in a fourfold process: (1) the 25 interviews were divided into two blocks of ten and one of eight; (2) the first two blocks were used to extend ideas and gather the most interesting topics of the interviews in an Excel spreadsheet (Appendix B) (3) the last set of eight interviews was used to validate the ideas and preliminary conclusions; (4) all interviews were brought together one last time and analysed.

The Excel table in Appendix B was developed in the following way: (1) the names of the participants are displayed horizontally; (2) the ideas these participants shared during the semi-structured interviews are displayed vertically. Appendix B also demonstrates what sub-questions these insights initially came from, as the semi-structured interviews were divided into three sets of four sub-questions. These four sub-questions were vital for the coding of the answers; refer to Appendix C for the sub-questions. The coding was conducted as follows: (1) the researcher highlighted the most interesting insights offered by the participants during the semi-structured interviews with a marker and added a comment to each one. These comments were narrowed down to 81 phrases, and listed at a later stage in the Excel table (Appendix B); (2) Once the Excel table (Appendix B) was created, these 81 phrases (Appendix C) were related back to the three main questions and divided into three phases (see Appendix D) .

4.6. DATA ANALYSIS

4.6.1. Narrative Analysis

The narrative analysis looked at the story's content and what it tells about the narrators to help reach the research goals of learning more about two questions: 1) How do Apache Hadoop creators and consumers generate innovation streams in their current markets? 2) How can innovation streams evolve from community base to enterprise platforms? A total of 25 semi-structured interviews with workers at 23 organisations in four Brazilian cities were undertaken. According to Creswell (2007), theoretical saturation can be obtained with 20 interviews. For Guest et al. (2006), saturation can be achieved with as few as 12 interviews. In the present study, no new codes occurred when coding the interviews in the fourfold process by the second block of ten interviews. This, in turn, indicated that saturation was achieved.

The participants were chosen to represent a wide range of perspectives. To find responders, we employed theoretical sampling. For example, participants from both big data and non-big data organisations in technical and non-technical roles were interviewed.

Theoretical sampling selects respondents who should vary in critical skills of interest. For example, we expect technicians to provide more technical jargon and essential technical processes that need to be translated to a managerial vocabulary and non-technical participants a more business and organisational vocabulary. Ten technical experts and ten non-technical experts were among the attendees. Only one of them was female, and they were all from Brazil. Participants varied from recent MBA graduates and those just starting their careers to experienced executives in their 40s and 50s. Participants held different positions in different firms using or selling Apache Hadoop products and services. See Section 5.7.1. Key professions for details on the various stakeholders' roles in Apache Hadoop projects.

4.6.2. Generalisation and Theoretical Saturation

Generalisation is widely accepted in quantitative research and somewhat controversial in qualitative studies. This is because qualitative studies aim not to make general statements about people's lives. Instead, qualitative research attempts to give a rich, contextualised understanding of a particular aspect of human experience by studying a variety of specific contexts (Polit and Beck, 2010).

Positivist (Popper, 1963) tends to back up their claims of knowledge with physical measurements or theories that have to be proven to be true across space and time, which is called universal. Interpretivists (Klein and Myers, 1999) believe that the interpretation of events must be general to apply to several contexts. As knowledge is constrained for systems theorists (Rescher, 1979), generalisation implies finding new bounds. Finally, according to argumentative epistemology (Crosswhite, 1996), a rational argument reveals its degree of generality, determining what kind of supporting evidence is suitable. In summary, each of these four epistemologies has a distinct set of goals for generalising. Still, these objectives are relevant to all of them (Metcalfe, 2005).

Walsham (1993) and Lee (1989) argue that the validity of conclusions made from one or more events is determined by the coherence and plausibility of the logical reasoning used to explain the outcomes of the events and conclude them rather than the statistical generalisability of the data. According to Walsham (1995), four generalisations are drawn from interpretative studies: idea development, theory creation, particular implications drawing, and the provision of rich insight. The critical point here is that theory plays a crucial role in interpretive research and clearly distinguishes it from just anecdotes (Klein and Myers, 1999).

Interpretative research reports will make generalising claims about the findings without commenting on how such generalisations are justified. A criticism of interpretive research is that generalisation is seldom mentioned, or it is applied only indirectly in the context of external validity, transferability, and confirmability (Ward-Schofield, 1993). Yet, generalisation seems to be unavoidable in interpretative research. Almost every study that is published will make some generalising assertion. Interpretive studies are often conducted at a micro-level, they require dense, detailed and contextualised description, and the micro-level detail of particular stakeholders is utilised to create a picture of that group as a whole (Williams, 2000).

4.6.3. Theoretical Saturation and Sample Size

A typical feature of assessing the quality of qualitative research is the need to determine the saturation of a data set, which indicates that the data include all the information required to address the research objectives. Apart from its broad importance, the idea of saturation is a source of contention among qualitative researchers (O'Reilly and Parker, 2012), who differentiate among the aspects of meaning, transparency, and applicability. When more observations and analysis show no novel themes, thematic saturation has occurred (Green and

Thorogood, 2004). It may be performed without fully understanding theme connections. Theoretical saturation happens when further evidence cannot improve the qualitative theory (Glaser and Strauss, 2009). Several approaches have been developed to measure saturation or handle the associated issue of predetermining adequate observations for qualitative research (Lowe et al., 2018).

Qualitative research has lately been criticised for its lack of rigour in sample size explanations (Marshall et al., 2013). Dworkin (2012) raised this subject, although it is seldom mentioned in the literature. Many reviewers are concerned about the sample size required for qualitative research results to be valid. When considering sample size in qualitative research, the idea of data saturation (Guest et al., 2006) is relevant. It is the point at which no new information or themes are seen in the data following the completion of more interviews or cases. The approach implies that a single case or interview is never satisfactory since data saturation can only be determined when at least two, and generally more, examples are reviewed. This concept of sampling until data saturation is achieved can justify its adoption (Boddy, 2016).

4.7. CONCLUSION

The chapter has discussed the methodology of the thesis. Based on the question design (‘how’ questions), the strategy was to carry out a piece of qualitative multiple-case study research. The study applied narrative inquiry to improve the theoretical variety and extend the theory of outbound OI and ambidexterity. For the data collection, theoretical sampling logic was used to select the participants from among the sampling population of Apache Hadoop professionals in Brazil. Primary (25 semi-structured interviews) and secondary (workshops, industrial reports and web magazines related to the Apache Hadoop) sources of evidence were used in combination to thematic coding techniques to analyse the data.

CHAPTER 5. APACHE HADOOP

5.1. INTRODUCTION

The meaning of “big data” is still the subject of some debate. It generally refers to data that is too big or too complex to process on a single machine. According to International Data Corporation’s annual Digital Universe study (IDC, 2014), the amount of data on our planet is set to reach 44 zettabytes (4.4 ×1022 bytes) by 2022, which would be ten times larger than it was in 2013. While no single entity is working with data of this magnitude, many industries are still generating data too large to be processed efficiently using traditional techniques (Landset et al., 2015). It can be said that we live in an age where data is growing by orders of magnitude faster than ever before. Data have become significant corporate assets and the foundation of new business opportunities. Organisations have radically altered the way they save, manage and monetise data. Businesses are acquiring data from different sources and in various formats. These data were previously considered to be worthless or too expensive to store (Zikopoulos and Eaton, 2011; White, 2012)

Big data describes pools of large datasets that cannot be analysed by exploiting conventional technologies and is the combination of diverse categories of data (e.g. traditional enterprise data, machine-generated sensor data and social data) (Dijcks, 2013). The difficulties associated with managing such vast amounts of data go beyond calculation and begin with the simple act of saving. Distributed computation is well established and vital in overcoming these difficulties and obtaining critical intelligence from large datasets (Polato et al., 2014).

Beulke (2011) has named five attributes of big data, the so-called 5Vs: data volume; velocity; variety; verification; and value. Firms are becoming more competitive through the development of procedures for acquiring, evaluating and decoding data (Davenport et al., 2012). To implement these procedures, businesses are undergoing organisational and cultural changes. Firms have to be more responsive to changes to deal with data more efficiently and align these data with the firm's IT infrastructures for service and product development (Brown et al., 2006; Liftman et al., 2006). Organisations working at big data level have harvested new assets that did not exist at such large scales until recently. These new resources upgrade existing products and services and/or create entirely new ones. One of the ecosystems that is capable of handling the amount of data cost-effectively and efficiently is the Apache Hadoop.

The next section discusses what the Apache Hadoop is. Section 5.3. summarises the importance of open source software for big data; Section 5.4. covers some of the components of the Apache Hadoop; 5.5. introduces some Apache Hadoop vendors, and 5.6. puts together the take-aways of the chapter. It is critical to point out that the Apache Hadoop is a highly technical ecosystem. Therefore, the content of the chapter has been drawn mostly from non-academic sources.

5.2. WHAT IS THE APACHE HADOOP?

The Apache Hadoop is a collection of tools available from the Apache Software Foundation. In its original and most uncomplicated form, it consists of the Hadoop Distributed File System (HDFS) for storing and MapReduce for processing. Table 33 illustrates the differences between the Apache Hadoop and older types of database technologies.

Table 33. The Differences between the Apache Hadoop and older types of Database Technologies

Data Base Technology	HDFS is rather a distributed file system.
The essences of relational databases are based on the ACID principle: atomicity, consistency, isolation and durability.	This is not the case with the Apache Hadoop. Thus, if an engineer wants to recreate a use case, he has to recreate all possible scenarios.
Cannot be easily scaled	One of the main qualities of the Apache Hadoop is the massive scaling of computing power and storage capacities at a relatively low cost compared to relational databases.
Although relational databases are also capable of handling unstructured data, when the amount of data gets large, performance becomes a problem.	The Apache Hadoop enables distributed processing of very large amounts of data
The Apache Hadoop's strength is distributed between processing and scalability - there is no limit to the size of the data.	Relational database queries are always slower than an SQL search in the Apache Hadoop

Source: Wuttke, 2020.

The Apache Hadoop is based on the Google File System (GFS) and Google MapReduce and developed by Dough Cutting. A major contributor from the very beginning has been Yahoo, and the organisation currently uses the Apache Hadoop extensively across its business (Dwivedi and Dubey, 2014). Table 29 pinpoints the differences between the Apache Hadoop and older types of database technologies (relational databases) based on Wuttke (2020).The Apache Hadoop is said to be the initiator of the big data era and can store very large amounts of data and process it at high speed. It was originally developed by Google Inc.

and has been distributed by the Apache Foundation since 2008. Today, the Apache Hadoop is the standard and the best-known big data ecosystem. The ecosystem consists of four layers and various components within the layers, Section, 5.4. - An Overview of the Apache Hadoop Architecture - breaks down and discusses the four layers and architecture the Apache Hadoop (Wuttke, 2008).

5.3. OVERVIEW OF THE APACHE HADOOP ARCHITECTURE

The enormous volume and variety of data have overwhelmed traditional ways and techniques used to manage and process data. Traditional methods have done an excellent job until now, but they are inflexible and costly. Apache Hadoop uses distributed affordable commodity hardware to manage, process and store vast amounts of data; in some cases, the networks consist of thousands of low-cost dedicated hardware computers within a single ecosystem. Key Features and Advantages of the Apache Cassandra. Table 34 summarises the The Apache Hadoop architecture.

Table 34. The Apache Hadoop Architecture

Layer	Function
Distributed Storage Layer	Every node in a cluster has its hard drive, memory bandwidth, and processing unit. The incoming data is divided into smaller data blocks and sent to Hadoop Distributed File System (HDFS) for storage. The HDFS stores three copies of every disk drive within the network. The HDFS master node saves the metadata for the individual data block and all its replicas.
Cluster Resource Management	Apache Hadoop manages the nodes to share resources between the application and users effectively. While MapReduce operates the resource management and data processing, YARN[4] splits the two functions. YARN allocates resources to different frameworks written for Hadoop.
Processing Framework Layer	The processing layer analyses and processes the incoming datasets into the cluster. Structured and unstructured datasets are mapped, shuffled, sorted, merged and reduced to smaller manageable data blocks. The same operations are dispersed across various nodes as close as possible to the servers with the data.
Application Programming Layer	With the introduction of YARN in the Apache Hadoop 2 the development of new processing tools and APIs[5] As big data solutions advance, technologically many different other applications appeared in the ecosystems. The new projects usually focus on streaming, user-friendly interfaces, programming languages, messaging, fail-overs, and security are all an intricate part of a comprehensive Hadoop ecosystem.

Source: Kaplarevic, 2020.

As section 5.5 - The Components of the Apache Hadoop - will discuss in more detail, at the core of the Apache Hadoop is the Hadoop Distributed File System (HDFS), YARN, and

4 YARN is an application-specif resource management component
5 According to Mulesoft,:“API is the acronym for Application Programming Interface, which is a software intermediary that allows two applications to talk to each other. Each time you use an app like Facebook, send an instant message, or check the weather on your phone, you’re using an API”.

MapReduce. HDFS is a set of protocols used to store large data sets, MapReduce processes the incoming data, and Yet Another Resource Negotiator (YARN) was developed to improve resource management and scheduling processes in a Hadoop cluster. YARN can be viewed as an interface that connects the HDFS and YARN with numerous other open-source Apache projects to complement the Apache Hadoop: Apache Hive, Impala, Pig, Sqoop, Spark, and Flume. Table 31 gives an overview of these Apache Hadoop projects based on IBM (2020). The Apache Hadoop is divided into four different layers (Kaplarevic, 2020). Table 30 summarises the four different layers in the Apache Hadoop Architecture based on Kaplarevic (2020). Table 35 summarises the advantages and disadvantages of Apache Hadoop.

Table 35. Other Important Apache Hadoop Projects

Apache Hive	It is a data warehouse programme that supports the reading, writing and managing large datasets in distributed storage using SQL.
Apache Impala	It is a native analytic database for the Apache Hadoop.
Apache Pig	It is a framework for the analysis of large datasets.
Apache Flume	It is a distributed, secure and available service for the effective processing , aggregation and distribution of vast volume of log data.

Source: IBM, 2020.

5.4. THE IMPORTANCE OF OPEN-SOURCE SOFTWARE FOR BIG DATA

Big firms making profits need performance guarantee and a high level of service. What a horror scenario would that be for a telecommunication giant if the organisation could not charge their clients for their service for a couple of hours due to a technical issue with the platform performing billing. Or, what does it mean for an online retail business to be offline for a whole morning in the peak month of December? The point made here is that technology can fail, and in such a case whose fault is it? The open-source community? The chief technology officer (CTO) for picking the wrong solution for the organisation? Why is open source still the first choice for firms handling big data? The open access to the source code helps trouble shooting there is no doubt. But this alone do not justify the dominance of free and open-source software for big data analysis. There are three other very important aspects that cannot be left aside. Firstly, there is no need to build everything from scratch – there are already very good, and most importantly, tested solutions available for software engineers to build upon. Secondly, we have witnessed a share investment in big data technologies for decades. Not only Silicon Valley giants chipped in; other massive multinationals have been very active in backing advancement in the Apache Hadoop – Intel, IBM and Siemens are only the tip of the iceberg. Third, firms are sometimes subject to be locked-in when opting for a propriety software solution. These software services or products may or may not be compatible with other technologies. It can cost firms millions of pounds to change to another vendor or similar technologies. This is not the case when committing oneself to a particular Apache Hadoop vendor. Obviously, each Apache Hadoop vendor has its strength which make important to choose the right one. However, it is relatively easy to cancel the contract with one and move to another vendors.

As big data grows, and as more people and companies become aware of its possibilities and benefits, the number of open-source data ecosystems will grow exponentially. The development of a single platform that addresses all the possible issues inherent to big data is too costly, if not impossible. As a result, the direction that is most widely pursued is for each firm to engage in building a network that is close to its unique interest, either internally or by calling to the community (Coimbra de Almeida and Bernardino, 2015). The biggest names (users and contributors) in big data use open source ecosytems. Many of them are part of the same Apache family. There are many explanations for the growth of these open-source systems; one of which is how many individuals in different fields can work together successfully. When businesses share their work and encourage others to participate, the effects are beyond the eyes to discover new gaps and new prospects. Big data technology owes a lot to major companies like Google and Facebook. They continuously put their data and money back into the community. Technology seems to be evolving very rapidly, but this is not an immediate process. If businesses had to try to solve data issues on their own, without feedback or support from open-source community, it would be a painfully long process (Guess, 2016).

5.5. THE COMPONENTS OF THE APACHE HADOOP

5.5.1 Hadoop Distributed File System (HDFS)

The Hadoop Distributed File System (HDFS) is a unique file system that is used in combination with MapReduce. It functions in the same way as a big data storage system. Its main feature is to divide the 'big' data within HDFS between multiple machines in a network to minimise the costs of the searches for data to be processed (Veetil and Gao, 2014). HDFS is the most used component of the Apache Hadoop ecosystem and is implemented on a variety of platforms. HDFS has its origin at Google, and the idea was to move the processing power to where the data was located as it made more sense from a network bandwidth and latency perspective – so that each computer could map or reduce the data locally. The central concept behind it is that the best architecture is the one with the fewest delays between a computer and its data (O'Reilly, 2017). Table 36 illustrates the advantages and disadvantages of HDFS (Wuttke, 2020).

Table 36. The Advantages and Disadvantages of HDFS

Advantages	Disadvantages
Very large files: hundreds of TB Scaled to thousands of standard servers Automatic distribution and replication Fail-safe	Physical locations of blocks are not transparent Limited optimization potential for higher services (like Hive etc.)

Source: Wuttke, 2020.

HDFS is the fundamental file system of the Apache Hadoop. It stores file system metadata and application data separately, in contrast to other distributed file systems. HDFS saves metadata on a dedicated server called the NameNode and application data at different servers called DataNodes. All the servers are linked and exchange information with each other (Shvachko et al., 2010).

5.5.2. MapReduce

MapReduce is used to tackle the difficulties associated with linking the power of large clusters of computers. MapReduce enables professionals to concentrate on manipulating the data set and processing the data in distributed ways. It is responsible for network communication and fault tolerance (Condie et al., 2010).

5.5.3. Apache Cassandra

Apache Cassandra is also a distributed open-source database that is usually called a "NoSQL database". Apache Cassandra originated at Facebook to be used in their 'Inbox' feature. It was released as an open-source project in 2008. It was developed to manage "big data" workloads by distributing data, reads and writes (eventually) across multiple computers with no single point of failure[6]. The Apache Cassandra is a peer-to-peer distributed system that consists of clusters of nodes that can read or write requests (Guerrero, 2018). Table 37 illustrates the key features and Advantages of the Apache Cassandra.

6 According to Riley (2018), single point of failure means: "A potential risk posed by a flaw in the design, implementation or configuration of a circuit or system in which one fault or malfunction causes an entire system to stop operating."

Table 37. The Key Features and Advantages of the Apache Cassandra

Open Source	Open-source software is well-accepted, and many organisations have adopted infrastructure for managing data as free and open-source software. Free and open-source software is characterised as having a higher speed of innovation and faster adoption than proprietary software in big data analysis.
Flexibility and Familiar Interface	The Cassandra Query Language (CQL) feels like SQL – data professionals have a fairly easy transition to big data.
High Performance	Most traditional databases are based upon primary/secondary architecture. While the primary architecture is capable of reading and writing operations, the secondary architecture is only qualified to perform read operations. It can result in latency, higher costs and lower availability at scale. The Apache Cassandra operates differently - all nodes can perform all read and write operations.
Active Everywhere & Zero-Downtime	As Apache Cassandra can perform read and write operations in every node, data is very rapidly copied across cloud environments and geographies. If a node fails, users are directed to the next nearest node. The redirecting happens automatically, and the users do not notice it. Thus, applications are always available, and data is still accessible and never lost.
Scalability	A traditional database is usually scaled vertically – adding new, more expensive machines. It is not only more costly but also a time-consuming process. The Apache Cassandra can scale-up horizontally by simply adding more nodes to the network.
Replication	As data volume increases for most businesses, they are not moving to multi-data centres, hybrid clouds or even multi-cloud deployments to take advantage of the strengths of a distributed data centre. Apache Cassandra is an economical way to handle data without getting locked into any single provider's ecosystem.

Source: DataStax, 2020.

Apache Cassandra was designed to take advantage of multiprocessor/multicore machines and to run across multiple data centres. It scales to hundreds of terabytes and can best perform in scenarios where a vast amount of data – big data – need to be managed efficiently. It was released as open-source in July 2008. Still, the codes were only controlled by Facebook engineers – the community around the Apache Cassandra was non-existent. Eventually, between 2009 and 2010, the project was released under the Apache Foundation. The community around Apache Cassandra started to crystallise; firms such as Twitter, LinkedIn, and Apple, as well as independent developers, began to contribute to the project (Carpenter and Hewitt, 2020). Table 33 summarises the key features and advantages of the Apache Cassandra (DataStax, 2020). DataStax gives users and enterprises the freedom to run data in any cloud at a global scale with zero downtime and zero lock-ins. More than 450 of the world's leading enterprises, including Capital One, Cisco, Delta Airlines, eBay, McDonald's, Sony and Walmart, use DataStax to build data architectures for real-world outcomes, (LinkedIn, 2020).

5.5.4. Apache Sparks

As in all other components of the Apache Hadoop, Apache Sparks is an open-source data-processing engine for large data sets. It is designed especially for streaming data, graph data, machine learning and artificial intelligence applications. It can operate up to 100x faster than the alternative – MapReduce. Data scientists favour Apache Sparks because it includes API programming languages such as Scala, Java, Python, and R. Spark was developed in 2009 at UC Berkeley and is now run by the Apache Software Foundation. Apache Sparks is one of the most popular components of the Apache Hadoop ecosystem and widely used by several commercial big data offerings (IBM, 2020). Table 38 compares Apache Sparks and the Hadoop Distributed File System.

Table 38. The Apache Sparks and Hadoop Distributed File System in Comparison

	Apache Sparks	HDFS
Processing	In-Memory / Hard Disk	(HDFS)
Speed	In-Memory up to 100 times faster than Apache Hadoop	MapReduce always reads and writes from the hard drive, so not so fast
Usage	Batch, streaming, machine learning and graph processing	Batch processing only
Real-Time Capability	Up to millions of entries per second	Not Applicable
Robustness	Error tolerance available – no restart in the event of errors	Error tolerance available – no restart in the event of errors
Security	Apache Spark still has room for improvement compared to the Apache Hadoop	More security than Apache Sparks

Source: Wuttke, 2020.

5.6. APACHE HADOOP VENDORS

5.6.1. Cloudera

Cloudera is the most popular vendor and it was the first to offer and deliver an enterprise solution for advanced data analytics and machine learning. It specialises in additional support and consulting services based around the Apache Hadoop. It was founded in 2008 by four Silicon Valley computer engineers: Christopher Bisciglia (a senior software engineer from Google), Amr Awadallah (vice-president of product intelligence engineering at

Yahoo), Mike Olson (ex-Oracle embedded technologies vice-president) and Jeff Hammerbacher (former Facebook data manager). Cloudera maintains its headquarters in Palo Alto, California, but has a global presence with more than 3000 employees across 24 international offices. It has joined forces with HortonWorks – another leading contender in the Hadoop big data space – in 2019. It can count on major key players in the technology world – Oracle, Dell, Microsoft and Hitachi. Cloudera has also made substantial progress in the artificial intelligence (AI) arena since the acquisition of Fast Forward Labs (a leading AI research and development organisation) in September 2017. The ultimate aim of the new investment is to become a significant player in the industrialisation of machine learning and artificial intelligence (AI) to provide practical solutions to global business problems. According to Forbes (2016), Cloudera is listed number 5 in private cloud companies across the globe. Since April 2017, its shares have been listed on the New York Stock Exchange (Parker, 2020).

5.6.2. DataStax

DataStax is the main contributor to Apache Cassandra and has so far contributed more than 85% of their permanent codes. DataStax was founded in 2010 by former Rackspace engineers Jonathan Ellis and Matthew Pfeil. But its history began in 2008, when Apache Cassandra was release by the Apache Software Foundation and Jonathan Ellis received an offer from a venture capitalist, John Vrionis at Lightspeed Venture Partners, to invest in new "big data" technologies, such as the Apache Hadoop and NoSQL databases for transactional workloads. Since 2013, DataStax has continuously added new functionality to the Apache Cassandra, and it eventually released its DataStax Enterprise, which is widely accepted in the community (DataStax, 2020).

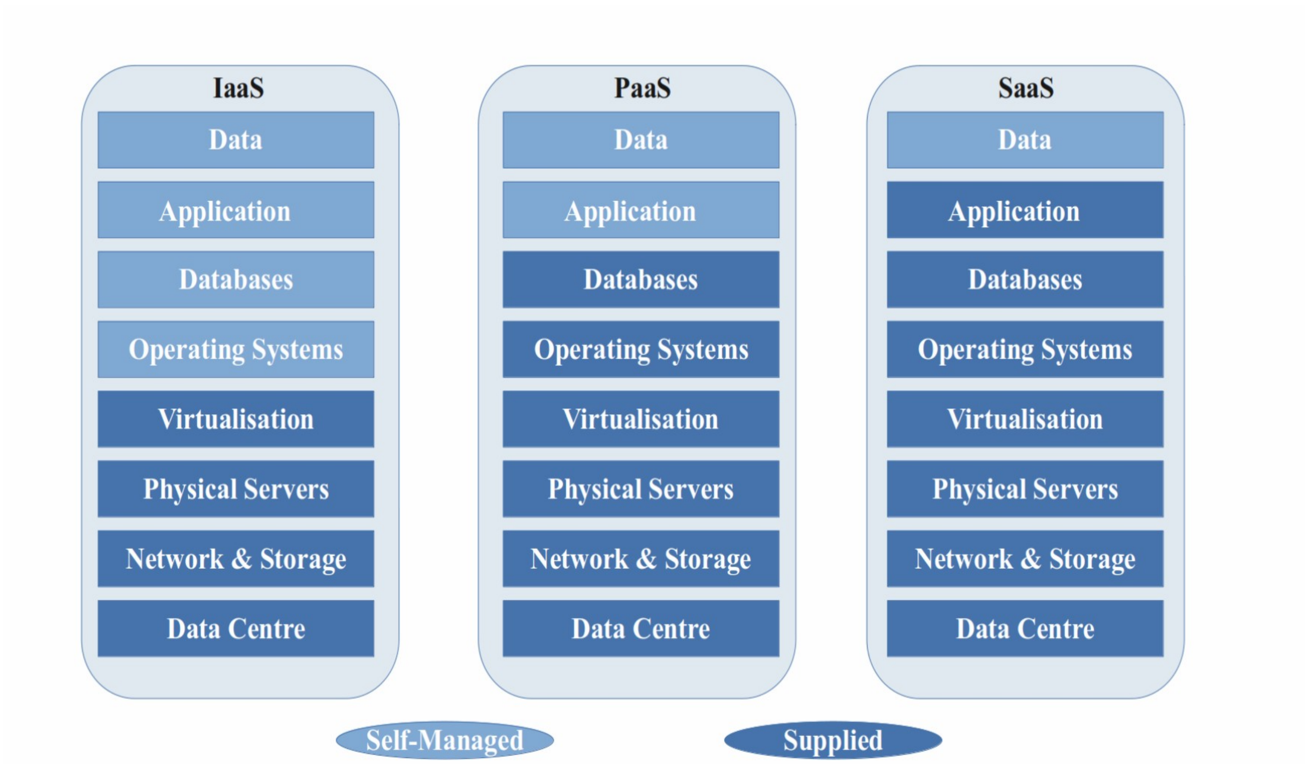
5.6.3. Hadoop as a Service

Cloud computing is the physical relocation of computation: e.g., on-demand computing, software as a service, or the Internet as a platform. The raw format can vary from a simple video streaming service, such as Netflix, to very complex systems, such as the Apache Hadoop. In practice, cloud computing is a computer system whereby some of the components (software and hardware) reside somewhere hidden across continents and unseen by the end-users. A substantial volume of computing activities is moving away from the desktop/laptop and office server rooms. The shift is having a significant impact on computational ecosystems across end-users – from the casual user to the software developer, the IT manager, and even the hardware manufacturer (Hayes, 2008).

Generally, firms are responsible for the whole stack, end to end, e.g., hardware, network, virtualisation, operating systems, middleware, software and data. Using traditional IT settings, IaaS, PaaS and SaaS are each on the way to allowing a layer of abstraction. IaaS takes the responsibility of end-users away from the physical computer, the network, storage, and the technology needed to virtualise the resources. PaaS goes further and abstracts away the management of the operating system and middleware. SaaS provides the entire end-user application as a service (IBM, 2019). The simplest way of thinking about the differences between the "X"aaS categories – IaaS, PaaS, and SaaS – is by identifying which component of the stack is run by whom. Each organisation has to critically judge its own IT capabilities and ask two questions: What elements do I manage? What components do I want the vendor to be responsible for? Figure 5 illustrates possible vendor abstractions.

As in 'X'aaS, Hadoop-as-a-Service (HaaS) is how Apache Hadoop vendors may share the responsibility of the IT settings for a fee. In comparison to classical infrastructures, HaaS offers the following advantages for businesses: (1) fast and flexible deployment without changes to the physical infrastructure; (2) scaling capabilities and multiple tenants in shared data centres. Typically, these are achieved by abstracting the physical hardware of a data centre with a cloud platform.

Figure 5. *Cloud Computing Vendor Abstractions*



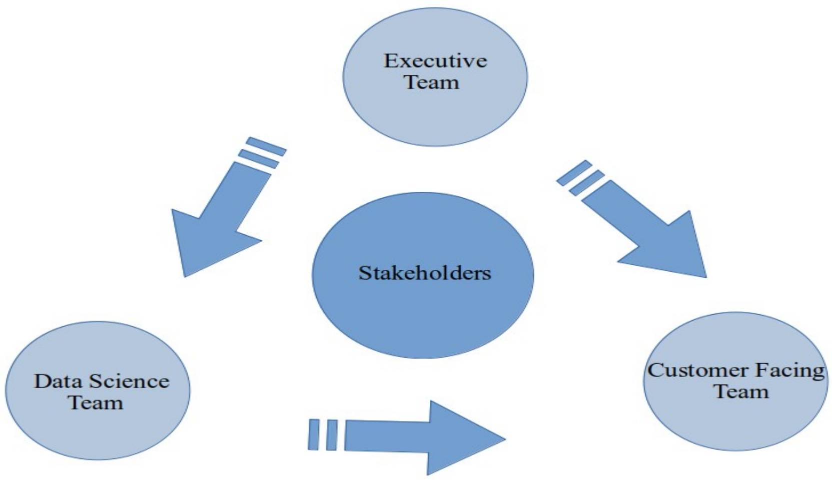
Source: IBM, 2009

5.7. ORGANISATIONAL AND STAKEHOLDER STRUCTURES

5.7.1. Key Professions

Three primary departments and four key professions are identified in generic Apache Hadoop projects: data scientists, software developers, solution architects and administration. The four professions are explained in more detail in this section based on upGrad (2021) and illustrated in Figure 6. The business administration part is not discussed in detail in this work as it does not differ much from other industries. Yet, it is worth mentioning that some executives or directors of Apache Hadoop service firms have a technical background.

Figure 6. *Departments within Apache Hadoop Firms*



A Data analyst processes data using a diversity of big data tools. Data analysts typically deal with numerous data types such as unstructured, semi-structured, and structured data. Data analysts master frameworks such as Apache Spark and Apache Hadoop to analyse the various forms of data. A data analyst's primary goal is to identify hidden potentials in data to assist organisations in increasing revenue and changing their strategy. Data analysts must have exceptional problem-solving and mathematics abilities. They study historical trends, generate patterns and produce reports on the same. Table 39 summarises the roles and responsibilities of data analysts .

Table 39. The Roles and Responsibilities of a Data Analysts

Roles and Responsibilities of a Data Analysts
To identify, analyse and interpret trends or patterns in large amounts of data
To possess the capacity to investigate and comprehend complex database architecture and table relations inside the database
To evaluate a huge number of data and develop a clear approach to a question or a new strategy
To understand how data structure requirements can help with data extraction from different sources

Source: upGrad, 2021.

The bedrock for big data analytics is big data architecture. It is the overall framework for managing vast volumes of data to be studied for business reasons, steering data analytics and providing an environment in which big data analytics tools can extract essential business insights from otherwise confusing data. In addition, the big data architectural framework acts as a design for big data infrastructures and solutions, logically outlining how solutions will operate, the components employed, how the information will flow and security considerations. Big data analytics architecture generally consists of four logical levels and conducts four critical operations (Omni Sci, 2021).

Consequently, the solution architect is responsible for analysing real-world problems and developing a plan to address them and solve difficulties, while utilising their knowledge and the capabilities of the big data framework. The solution architect is in charge of all tools, including programming language selections. Therefore, a solution architect must have a thorough understanding of the various big data technologies and knowledge of frameworks, license fees, and alternative open-source tools (upGrad, 2021). Table 40 summarises the roles and responsibilities of a big Data solution architects

Table 40. Roles and Responsibilities of Big Data Solution Architects

Roles and Responsibilities of a Big Data Solution Architects
To be well-versed in the design of large data architectures
To have expertise in real-time and batch-based big data processing, business intelligence, and machine learning
To maintain storage infrastructure and migrate data as needed

Source: upGrad, 2021.

The programmer is a vital member of the team. He creates all of the code that conducts repetitive and conditional actions on data sets. Shell Scripting, Java, and Python are the most prevalent languages that programmers work with. The code that programmers write is kept in a database or flat files; thus, a fundamental knowledge of the file and database systems is required. A solid understanding of the various programming languages, mathematical and statistical abilities and an analytical thought process to produce excellent code are essential (upGrad, 2021). Table 41 summarises the roles and responsabilities of a programmers.

Table 41. The Roles and Responsibilities of Programmers

Roles and Responsibilities of a Programmers
To develop and maintain Advanced Dashboard on Real-Time (Python)
To create interface with third-party systmens for data
To ensure the functionalities supplied are both reliable and available.
To define, create, and launch new features, collaborate with cross-functional teams.
To fix bugs and increase the performance of the program are priorities.

Source: upGrad, 2021.

Finally, there is the administrator, who is in charge of the complete infrastructure and oversees all of the data and the tool of the ecosystem. The job entails looking after all of the nodes in their network setups. administrators must ensure that sufficient time is available to support all important data activities. The Admin is responsible for the installation of various tools and the management of hardware clusters. A thorough grasp of the file system, operating system, computer hardware, and networking are among the skills that an administrator must possess (upGrad, 2021). Table 42 summarises the roles and responsabilities of administrators.

Table 42. Roles and Responsibilities of Administrators

The Roles and Responsibilities of Administrators
To manage and configurate all tools
To manage all application clusters in big data infrastructure
To identify application vulnerabilities and coordinate mitigating efforts.

Source: upGrad, 2021.

5.7.2. Apache Hadoop Market Segment and Key Stakeholders

Components, deployment models, organization size, and end-users are all used to segment the Apache Hadoop market. It is divided into three deployment types: on-premise, cloud, and hybrid. It is divided into two types of organizations based on their size: SMEs and large organizations. The end-users are manufacturing, banking, financial services and insurance (BSFI), retail and consumer goods, IT and communications, health care, government and defence, media and entertainment, energy and utilities, and commerce and transportation. Key market players include Amazon Web Services (AWS), Cisco Systems, Cloudera, Datameer, Dell, Google, IBM and MapR (AMR, 2021). Table 43 compares the Apache Sparks and Hadoop Distributed File System.

Table 43. The Apache Sparks and Hadoop Distributed File System in Comparison

KEY MARKET SEGMENTS			
COMPONENT	**DEPLOYMENT MODEL**	**ORGANISATION SIZE**	**END USER**
Hardware Software Services	On-premise Cloud Hybrid	Large Enterprise SMEs	Manufacturing; BFSI; Retail & Consumer Goods; IT & Telecommunication; Healthcare; Government & Defense; Media & Entertainment; Energy & Utility; Trade & Transportation

Source: upGrad, 2021.

5.8. THE TECHNICAL STTRUCTURES

The chapter has sofar dicussed different aspects of the organisational and stakeholder structure of the Apache Hadoop ecosystem. The next section focus on the technical structure, the particularities of different industry sectors and how firms utilise the Apache Hadoop. Additionlly, it introduces global and domestic background of the different sectors.

5.8.1. Telecommunication Sector

5.8.1.1. Global Background

Data is omnipresent in the telecommunications sector. Every mobile phone and computer and, most recently, television in our living room at home, is connected to the Internet. At the same time, the telecommunications sector is ubiquitous for most businesses today. Modern organisations rely on information and communication technology (ICT) and the Internet to process transactions every second. Aside from human-to-human communication and human-to-machine communication, nowadays machines are also talking to each other. According to a projection by Deloitte, there will be about 32 billion devices generating 44 trillion GB of data by the end of 2020. Machine-to-machine (M2M) communication is forecast to surpass the amount of data produced by humans by 2022 (Deloitte, 2015). On a global scale, the sector is predicted to grow to over £76 billion in revenue for service providers (McKinsey Global Institute, 2016). According to Forbes[7] (2018), the telecommunications sector leads the adoption of big data. Dresner Advisory Service estimates that 95% of telecommunications firms globally have already adopted big data technologies for businesses. Mobile data services dominate the sector and there are expectations of extended growth for the next-generation access (NGA) technologies, e.g. 5G mobile generation.

In addition to the growth of data and devices connected to the Internet, the sector has experienced a considerable structural shift over the last two decades. Technological innovations and adjustments in regulatory processes around the world have had a big impact on the sector and connectivity.

5.8.1.2. Domestic Background

Brazil, like many other emerging and developing countries, has followed the same pattern as other modern economies. With the government monopoly of the sector coming to an end and privatisation becoming a reality, Brazil has adhered to similar privatisation models to its international counterparts.

Initially, there was competition for market share between players around GSM and 3G mobile telephony and data. Later, the race intensified when 4G and 5G mobile technology came into play. According to a report conducted by McKinsey (2016), in Brazil today there are 241 million mobile phones, 41 million fixed lines and 27.5 million households with a

7 https://www.forbes.com/sites/louiscolumbus/2018/12/23/big-data-analytics-adoption-soared-in-the-enterprise-in-2018/#1b876736332f

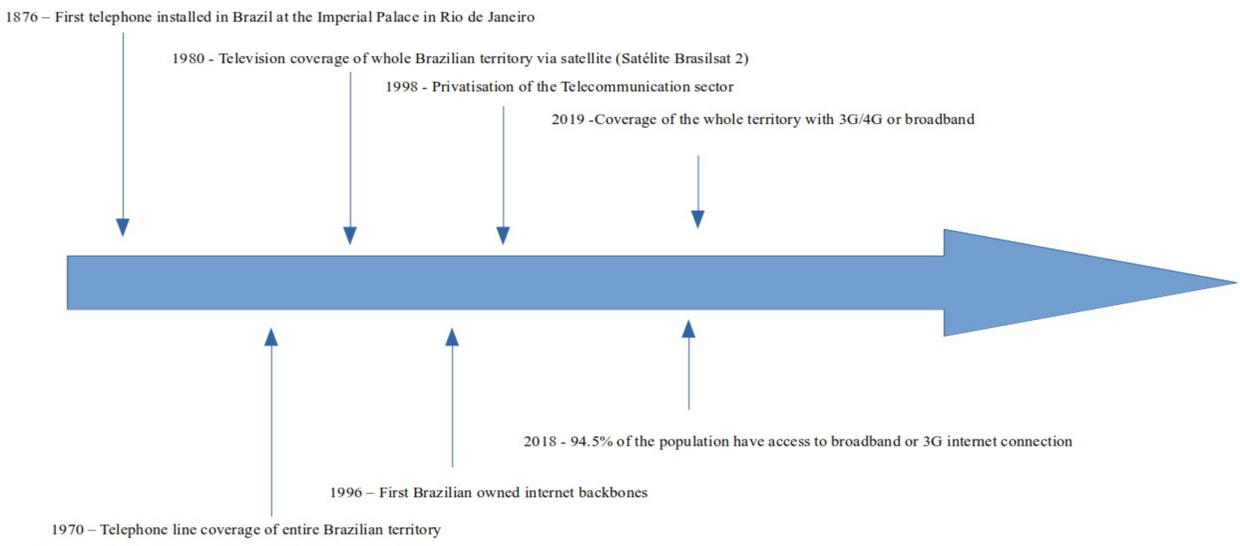
broadband connection. Since privatisation in 1998, there has been £123 millions of investment in the sector, putting Brazil on the map of connected societies. It has elevated the nation's telecommunications infrastructure to the fifth largest in the world. The value created by the telecommunications sector in Brazil in 2016 was around £1.7 billion. The privatisation of the sector has contributed to the wide diffusion of connectivity in the country. Most of the population did not have access to basic telephone services before privatisation (Telebrasil, 2018).

The privatisation of the Telebrás System took place on the 29th of July 1998 through 12 auctions on the Rio de Janeiro Stock Exchange (BVRJ) for the sale and control rights of three fixed-line, one long-distance and eight cell-phone holdings. It was the largest privatisation operation ever carried out in the world. Thirty-one licences were given out, and the government raised a total of £16 billion, 63% over the expected price (Jurado da Silva, 2015). By June 2018, 4,071 municipalities out of 5,570, the equivalent of 94.2% of the population, had access to broadband or 3G connectivity at their disposal. By the end of 2019, the headquarters of all municipalities in the country were connected to a 3G/4G or broadband network. The mandatory numbers were established by the Agência Nacional de Telecomunicações (Anatel).

Over the last two decades, numerous mergers and takeovers have taken place in the sector in Brazil, resulting in a state of limited competition in which the market is shared by a small number of service providers. The decrease from 31 to nine operators in 2006 has reinforced the oligopoly[8] market structure. Figure 07 shows an historical timeline of the telecommunications sector in Brazil.

8 A state of limited competition, in which a market is shared by a small number of service providers.

Figure 07. *Historical timeline of the telecommunications sector in Brazil*



Source: Jurado da Silva (2015).

The big winners of the privatisation of the sector were VIVO, TIM and Claro Group. These three organisations share the domestic market today and provide connectivity for over 250 million Brazilians. Figure 6 illustrates a historical timeline of the most important events in the sector – from the first telephone installed at the Imperial Palace in Rio de Janeiro to the coverage of the entire Brazilian territory with 3G/4G or Internet broadband.

5.8.1.3. The Technology

The sector is very diverse and consists of a wide range from traditional voice and text messaging business models to mobile data services (3G, 4G and 5G) and Internet services (fibre and broadband). Firms that once relied upon voice and messaging technologies are now driven by data instead. Table 44 summarises the data structures and formats in big data.

Table 44. Data Structures and Formats in Big Data

Type	Formats	Usage
Structured	Conventional datasets	Common data used to describe a page. This data type is used by search engines such as Google.
Semi-structured	XML[9], JSON[10]	The data consists of metadata to group information and specify how it is stored.
Unstructured	Logs, chat rooms, audio and videos	This can include everything from emails to audio or video files and presentations or business documents.

Source: PWC, 2018.

Most recently, telecommunications firms are also competing with broadcasters in the distribution of audiovisual content via Internet streaming. Table 45 summsrises the data in the telecommunication sector big data.

9 The language is widely used in the World Wide Web and is applied to encode documents so that humans and computers can make sense of them.
10 JavaScript Object Notation – The language used to stored data structures in JavaScript.

Table 45. The Data in the Telecommunication Sector Big Data

Type of data	Segments	Explanation
Call detail data	- Duration of call - Origin/destination - Call period - Area code	Call detail records (CDRs) consist of data generated by customer telephone calls and other telecommunications transactions, e.g., text messages.
Network data	- Configuration of equipment - Error logs - Network management	This can be understood as the data shared by computers and other devices in a network, e.g., signal towers. Operators use the data to overview their technological infrastructures.
Customer data	- Data generated by customers - Personal customer data	Operators hold this data about their clients. It consists of personal data, such as name or gender. Additionally, it includes the client profile data. This includes the type of device and other services the client uses or has applied for or used in the past.

Source: PWC, 2018.

The industry gathers massive amounts of data. The Apache Hadoop assists telecommunications organisations with finding hidden patterns and correlations in the data sets the firms collect from customers or from data the telecommunications firms create themselves. Table 46 summarises the key Applications for Big Data in the telecommunications sector

Table 46. Key Applications for Big Data in the Telecommunications Sector Based on Tata Tele

Categories	Applications	Details
Customer experience enhancement	Targeted marketing	The personalisation of a product offering, e.g., top-up data plans, roaming and international calls. The analyses are based on billing data, support requests, demographic data and purchase data.
	Prediction of customer churn	The operators can predict when customers may want to end their relationship or not extend their existing contract. Telecommunications firms use data such as call details, network performance and social media to anticipate and limit problems affecting customers.
	Journey analytics	This is the data that shows the interaction between customers and organisations at various stages during their relationship. It helps telecommunications operators to fine-tune their new offers and campaigns.
	Proactive customer care	To predict issues and fix them before customers are affected.
Network Optimisation	Planning and optimising the capacity of the network	This data is used by telecommunications operators to overview network usage and user density among other things such as traffic and location data. It helps them to

		allocate resources more appropriately due to better monitoring of the infrastructure. It is used to manage and forecast network capacity.
	Network investment planning	Operators use the data to better allocate their future investments. Decisions are based on future connectivity needs by forecasting traffic in particular regions or the need to improve customers' experiences.
	Real-time analytics	These are often collected from cell towers. Operators can react to events occurring in real time. They can use the data to proactively react to failures or outages.
Operational Analytics	Cyber and data security	The data is used to identify in advance cyber threats and anomalies when data is accessed. It is directly to do with the complexity and growth of data in businesses in recent times.
	Customer care optimisation	This looks for the best way to reach customers, e.g. through platforms or campaigns. It enables firms to strategically use multiple campaigns across platforms. It assists firms by limiting contact fatigue.
Monetisation	Data analysis as a service	Some applications use location information to provide a range of services to users. Telecommunications operators use location information and data showing customers' preferences and offer this data to third parties.
	Data monetisation	One of the most popular usages of this is

	ride-sharing by Uber. Depending on the location of users, data are sold to third parties.

Source: PWC, 2018.

The combination of the segments has resulted in a considerable amount of data being generated. Table 1 summarises data structure and formats in big data. Table 3 illustrates the types of data, the different segments and how they are used. There are four key big data applications in the telecommunications sector: (1) Customer experience enhancement, (2) Network optimisation, (3) Operational analytics, and (4) Monetisation. Table 3 summarises the critical applications of big data in the telecommunication sector.

5.8.2. Financial Sector

5.8.2.1. Global Background

Big data technology is ubiquitous in the financial sector and banking. Firms can benefit strategically from big data in the generation of new revenue streams through personalised recommendations, or in the process of orders and the improvement of security. The sector has become increasingly diversified, putting financial institutions under pressure to get to know their customers (KYC) and themselves better, and data plays a significant role in the process.

A recent Interactive Data Corporation (IDC) report has forecasted that the global earnings for big data and business analytics will exceed $203 billion in 2020. The same report judges the banking industry as one of the top drivers of growth. The expansion can be associated with the eradication of friction for customers, e.g., online payment platforms and peer-to-peer lending. Although technology has introduced some changes in the sector, some more traditional financial institutions have yet to go through technological change and innovation (Chen, 2020).

5.8.2.2. International Agreements and Standards

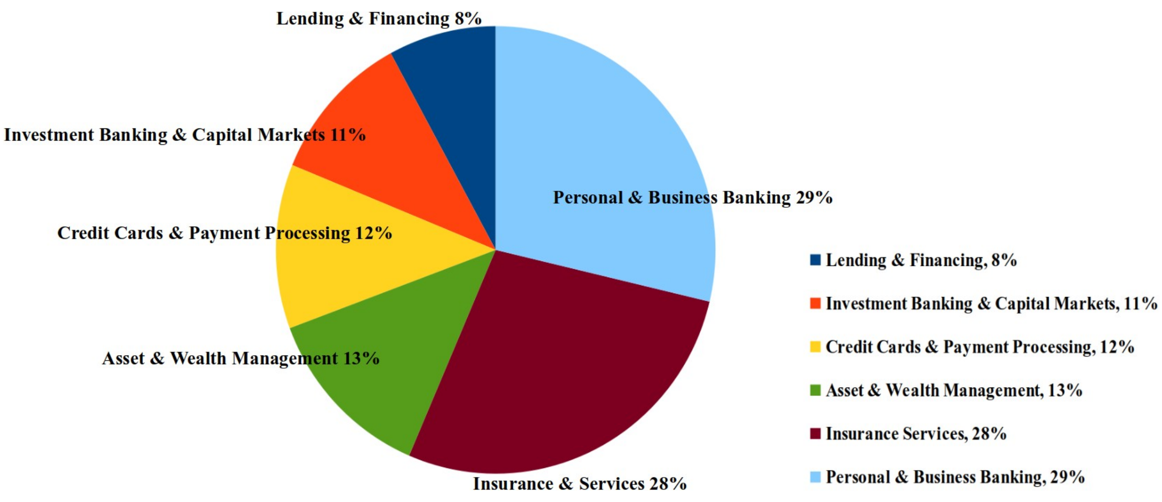
The Basel Committee is one of the authorities in addition to domestic juridical representatives that regulates and supervises financial institutions. Formerly known as the Committee on Banking Regulations and Supervisory Practices, it was founded in Switzerland by a group of ten countries at the end of 1974 following the turmoil in international currency and banking markets. The committee was set up to improve financial security by bettering the quality of bank supervision. It acts as a forum for mutual effort between its members' countries on banking supervisory matters, and Brazil is one of the 28 affiliated jurisdictions. Since the mid-1970s, the Basel Committee has published three documents commonly known as Basel I, Basel II and, most recently, Basel III. Basel III was written in response to the financial crisis of 2007–09 (BIS, 2020).

Banks and other financial institutions not only obey national legislation but are also under strict regulation by international bodies. Such international agencies advise their members on issues such as capital, market and operational risks. Of interest here are the operational risks associated with IT infrastructures published by the Basel Committee on Banking Supervision in 2003 under the document of BCBS 23923.

The importance of BCBS 23923 lies in the requirements for technological infrastructures in financial institutions. It calls for financial institutions to lay out, develop and cultivate data architectures and IT infrastructures that are fully risk-assessed at all times. In more detail, this document regulates (1) the risk of data aggregation capabilities; (2) the institutional creation of data taxonomies and architectures across firms, and (3) the duties and accountabilities of business and technology owners. It specifies the responsibilities of *Apache Hadoop* vendors when they perform services for financial institutions (BIS, 2020).

In addition to Basel III, Brazil also follows the most recent European Union recommendations on data privacy released under the General Data Protection Regulation (GDPR). Brazil's Lei Geral de Proteção de Dados (LGPD) delivers clarification to the Brazilian legal framework. It outlines the different statutes that presently dictate personal data (online and offline) (GDPR, 2020). It is imperative to consider the existence of the two documents to understand why compliance is required in some organisations. The documents are not of further relevance and will not be discussed in more detail in this study.

Figure 08. *Investment in Big Data by the Different Segments in the Sector*



Source: Abraham et al., 2019.

Like other sectors of the economy, big data technologies have changed the way financial institutions manage their own and their clients' data. It is common in the sector to mine data from social media profiles or public records to enhance their information on clients. Additionally, they can look for information in news, tweets, blog posts and other online publications, and follow market trends and predict market upturns and downturns. Firms in all segments of the sector have already adopted big data technologies (Abraham et al., 2019). Figure 8 shows the diversity of the sector.

5.8.2.3. Domestic Background

Brazil is the largest and most modern financial market in Latin America. Since the 1980s, domestic banks have survived many episodes of hyperinflation. Most recently, Brazilian banks have experienced a more stable financial environment and appeared rather resilient to the most recent financial crisis in 2008. Before the 1994 Real Plan[11], domestic banks competed for deposits, and, with the capital gain, banks invested in inflation-protected government securities. The so-called "float"[12] enabled banks to expand their businesses by opening new branches and offering free banking services. This kind of strategy would not be profitable in low inflation environments. Nevertheless, the model facilitates some technological progress, and is specially designed to speed up financial transactions (Goldfajn et al., 2003).

Today, the system is governed by large financial conglomerates and is focused on short-term liquid instruments. The sector is led by three private and three public banks, and together they own about half of the sector. The Brazilian economy has been relatively stable over the last decade, despite the domestic and international recession, and the system has remained consistent for the most part. Although strongly influenced by the liberal reforms inspired by Thatcher and Bush, going back to the early 1990s, the government still plays a

11 Plano Real (Real Plan) was put into motion under very idiosyncratic circumstances without the support of International Monetary Fund in 1994. It overcame the deep-rooted high inflation with a home-grown financial reform programme (Bacha, 2003).

12 Generally known as inflation profits

dominant hand. The combination of privatisation versus state-owned institutions has resulted in a nearly 200% growth in GDP over the last 30 years; this is in line with the growth of other emergent economies (IMF, 2018).

To sum up, Brazil has witnessed two very unique financial systems over the last three decades. There was a gigantic shift from a state-owned monopoly system poisoned by hyperinflation before 1994 to the system we see today – a system characterised by a small number of public banks and a high number of foreign banks that do not profit from high inflation (Goldfajn et al., 2003). Figure 3 illustrates a simplified historical timeline of the Brazilian financial sector.

5.8.2.4. The Technology

Financial institutions must obtain and accumulate meaningful data to determine customers' repayment ability as meticulously as possible. The data from lenders are as crucial as that from borrowers. The data are usually described as hard data (credit history, income, employment, education level, tax records and financial statements) and soft data (opinions from loan officers, internal discussions and economic prospects) (Abraham, 2019).

The two types of data, in addition to the data generated internally, are usually used in financial risk modelling, risk analytics, diagnostic analytics, customer analytics, control monitoring and transaction data analytics. Table 47 illustrates the different categories and the applications of data in the sector.

Table 47. Application of Big Data in the Financial Sector

Category	Application
Financial Risk Modelling	Used for risk management, valuation and financial/regulatory reporting purposes.
Risk Analytics	Used to determine and react to risks and to answer to regulatory requirements. It helps to reinforce firms' analytics and risk management.
Diagnostic Analytics	Used to help organisations interpret data and gain valuable insights for decision-making, risk and compliance, resulting in the acceleration of growth.
Customer Analytics	Drives new growth and improves customer engagement. It facilitates customer-centric business transformation.
Continuous Control Monitoring	Helps to identify irregularities in transactions by using predetermined risk scenarios for detection of fraud or errors.
Transaction Data Analytics	Helps firms to process and navigate through data sourcing, integration and complex data calculations. It assists firms in solving internal business problems.

Source: PWC, 2018.

Financial institutions have the monopoly of holding money from third parties for investments and savings. Accordingly, financial institutions are required by law to collect data and hand it over to regulators, supervisors and other institutions, such as credit bureaus. The financial sector is firmly regulated and supervised (Abraham, 2019).

5.8.3. Retail Sector and Marketing

5.8.3.1. Global Background

The case studies cover the retail and marketing industries. They were joined due to the similarities of how the Apache Hadoop is used. The primary industry in this section is retail. Retail describes the trading of goods and services to consumers in relatively small quantities as opposed to wholesale. Retail is a reference to the old French 'retaillier' and it was originally a word used as a standard measurement for fabrics. Exchanges between sellers and buyers take place over different channels across a variety of industries – from physical retail stores to e-commerce. The global retail market size was approximately US$ 24 trillion in 2018 (O'Connell). The second sector discussed in the case study is marketing. Marketing has witnessed a slight and steady increase over recent years and was predicted to outperform by US$ 560 billion by 2019. The growth can be linked to the usage of technologies such as big data (SRD[13], 2019).

5.8.3.2. Domestic Background

The domestic retail sector consists of 1.6 million commercial companies and has secured around US$ 570 billion in revenue. The sector employs 10.7 million Brazilians and is responsible for most of the employed persons in the trade (IBGE[14], 2020). Brazil leads the marketing sector in Latin America and is one of the largest in the world. It was estimated to be worth nearly US$12.5 billion in 2019 (SRD[15], 2019).

13 Statista Research Department (https://www.statista.com/topics/2151/advertising-industry-in-brazil/)
14 Brazilian Institute of Geography and Statistics
15 Statista Research Department (https://www.statista.com/topics/2151/advertising-industry-in-brazil/)

5.8.3.3. The Technology

Big data is used in all phases of retail operations. Apache Hadoop can help a firm to find out what the most popular products are or to anticipate trends and predict where the demand will be for the articles. Once firms understand trends and needs, Apache Hadoop helps to optimise prices and locate customers who would be interested in the offers. Additionally, the Apache Hadoop can assist firms in approaching the potential buyers and in finally closing the deals. Most importantly, it can help firms to find out what to sell to customers next (Marr, 2015). Table 48 illustrates the key applications for big data in retail, based on Hitchcock (2018).

Table 48. Key Applications for Big Data in Retail

Categories	Applications	Motivation
Customer Behaviour	Big data can be used to improve customer conversion rates via personalised campaigns. It may help to increase revenue, predicting and avoiding customer churn and lowering customer acquisition costs.	Consumers dialogue through multiple interaction points — mobile, social media, stores, e-commerce sites and more. Big data helps firms to overcome the complexity and variety of data types that have to be managed.
In-Store Experience	Big data helps organisations to improve sales tactics by personalising the in-store experience. The Apache Hadoop assists firms in analysing store behaviour and measuring the effectiveness of the efforts made to push sales.	Big data offers new methods for product placement in-store.
Predictive Analytics and Targeted Promotions	Customers interact with the service provider through multiple channels, such as social media networks and web pages. Historically, the data gathered was limited to demographic data.	Service providers can convert the data generated by customers' transactions into valuable information.
Customer Journey Analytics	A service provider can use different platforms to access customers' locations through mobile location data, social media and e-commerce.	Customers shop around on the web and social media and pick the best option. A service provider can analyse all the data adequately and offer service along the customers' journey.
Operational Analytics and Supply Chain Analysis	Information about trends, patterns or deviations can improve decisions and better operational performances.	The insights can help service providers save millions of dollars and increase operational efficiency.

Source: Hitchcock, 2018

Traditionally, marketing campaigns' decisions were made by intuition and experience. Today, firms can rely on their data to make the necessary strategic decisions. Big data can offer marketing agencies tools to gain insights into what their customers require and desire (Saran, 2018). Table 49 illustrates the key applications for big data in marketing, based on Saram (2018).

Table 49. Key Applications for Big Data in Marketing

Categories	Applications	Motivation
Better Decision Making	The Apache Hadoop can help the service provider to better target the essential demands of customers by creating useful content. The service provider gathers data as customers browse the internet.	The service provider can develop different profiles of the buyer by using customer behaviour data, purchasing patterns, their preferences and their background.
Better Pricing Decisions	Apache Hadoop automates the prices of services and products. It saves time in price settings and improves pricing decisions.	Service providers can now set the prices driven by big data.
1.6 Recommendation of Web Content	The Apache Hadoop helps service providers with creating personalised strategy by using deductive and inductive research and how customer select their services and product.	The service provider can customise the contents they offer to customers by finding out which content will be more appealing to each customer.

Source: Saram, 2018

5.8.4. Public and Steel Industry

5.8.4.1. Global And Domestic Background

The Apache Hadoop is used across industries, including the public sector. Big data are used in the public sector for advanced analytics through automated algorithms, for improvements in effectiveness, and to improve internal transparency and to personalise services (Munné, 2016). In comparison to modern western economies, Brazil is one of the countries with fewer civil servants in relation to the total number of workers. Recent data indicate that in Brazil, around 12% of labour works for the government (Gala, 2020).

Steel plays an important role in today's world. It is one of the most important materials for construction and manufacturing. It is used for innovative processes across industries globally (World Steel Association, 2020). Brazil is one of the main steel producers worldwide. The production of steel in Brazil spans over 100 years, and it was initially subsidised to a large extent by the Brazilian government. Today, steel is one of the most important domestic export products and it is considered as one of the most important materials for overall economic growth in Brazil (InvestinBrazil, 2013).

5.8.4.2. *The Technology*

Big data and the Apache Hadoopcan help public sector organisations to make better decisions, which may result in service improvement narrowing the gap between firms in the private and public sector (Deloitte, 2020). The manufacturing industry is very complex due to the variety and depth of products that are manufactured. The industry is organised by the International Standard of Industrial Classification (ISIC) and is divided into two major segments: discrete manufacturing and process manufacturing. The industry includes sectors such as automobile, hi-tech, aerospace, chemicals, pharmaceuticals and metals (Datta, 2018). Table 50 illustrates the Key Applications of Big Data in the Public Sector (Deloitte, 2020).

Table 50. Key Applications for Big Data in the Public Sector

Categories	Applications	Motivation
Open Data	APIs (Application Programming Interface) Open Standards Open License	Data enables organisations to make better decisions. It helps organisations to save costs and to get to know their customers better. Open data in combination with open platforms develops the data's latent potential.
Workforce Analytics	Workforce analytics help to personalise sales by analysing customers' behaviour in-store and by measuring and testing its efficacy	Global economic decline has put the government's departments under pressure to improve the processes of planning and management of its workforce. Big data helps governments to reduce costs of core services for the general public and offers new approaches to handle the turnover of employees.
Finance Analytics	Finance analytics help the different departments of governments to enforce government regulation and compliance requirements Big data can help to reduce the budget and drive transparency with value for money. Big data can improve the ability of the public sector to compete with the private sector in some key areas.	The public sector is facing remarkable challenges. The Apache Hadoop can help the different departments of the government overcome some of that pressure. The Apache Hadoop offers innovative and real-time solutions for governments' financial reporting and management difficulties.

1.7	Customer Analytics	Customer analytics help governments to reduce budgets and help the public sector to compete with the private sector. Big data helps governments to offer more transparency about the service they offer.	Citizens have a better overall view of the market by shopping around on social media and the internet. The private sector is increasingly competing with the public sector in key services.
1.8	Risk & Regulatory Analytics	The Apache Hadoop helps governments to prevent fraud and corruption	As the public trust in government agencies decrease due to the track record of corruption among government members, the Apache Hadoop can help regulatory agencies to work more efficiently.

Source: Delloitte, 2020

The industry is organised by the International Standard of Industrial Classification (ISIC) and is divided into two major segments: discrete manufacturing and process manufacturing. The industry includes sectors such as automobile, hi-tech, aerospace, chemicals, pharmaceuticals and metals (Datta, 2018). Table 46 illustrates the Key Applications of Big Data in the Public Sector (Deloitte, 2020) and Table 51 illustrates Key Applications of Big Data in Manufacturing (Datta, 2018).

Table 51. Key Applications of Big Data in Manufacturing

Categories	Applications	Motivation
Optimize production and enhance efficiency	Process manufacturing Machine logs	The Apache Hadoop can help determine specific weak links within production processes.
Predict machine failure and reduce downtime	Big data is used for equipment breakdown prediction and to schedule maintenance.	The Apache Hadoop can reduce breakdown and previewing downtimes.
Optimize the supply chain	Big data helps to organise supply chain data from different sources and formats.	The global supply chains of prime material for manufacturers are becoming increasingly complex. Manufacturers are using the Apache Hadoop to overcome this type of difficulty and improve service and product delivery to the end consumers and partners.
Enhance product quality and cut manufacturing cost	Big data helps firms to improve products and services with fewer costs.	Big data can help improve efficiency by reducing testing costs of manufacturing plants and processes.

Source: Datta, 2018

5.9. CONCLUSION

Data have become a critical asset for innovative firms, and one way of handling the massive amount of data available today cost-effectively and efficiently is the Apache Hadoop ecosystem. This chapter has summarised the differences between the Apache Hadoop and older types of database technologies and offered an overview of the advantages and disadvantages of the system. Additionally, it has outlined the importance of open-source software for big data. The Apache Hadoop has been broken down into layers of architecture. Finally, some of the Apache Hadoop's components and vendors have been briefly summarised.

CHAPTER 6. RESULTS

6.1. INTRODUCTION

The last chapter discussed the Apache Hadoop in detail from the ecosystem's technical and organisational structures. This chapter introduces the four case stories. The case histories are based on 25 professionals from 23 organisations. There are four start-ups, nine domestic firms and ten multinationals. The 25 professionals share seven professions between them and each contributes to one or more case studies.

6.2. TELECOMMUNICATION CASE HISTORIES

6.2.1. A Truly International Sector

Telecommunications is a truly international sector. The Latin American and Brazilian markets reflect this reality. The Brazilian example demonstrates a domestic market that is dominated by three companies only: Vivo, TIM and Claro are the main providers and are all Brazilian subsidiaries of international telecommunications firms. Vivo, the largest telecommunications organisation in Brazil, is a subsidiary of Spanish Telefónica; TIM Brasil is a subsidiary of Telecom Italy, and Claro Brasil is a subsidiary of Mexican Móvil. These examples explain why there is a minimal technological discrepancy between firms using the Apache Hadoop in Brazil, Europe or the USA.

It can be said that most data problems within telecommunications firms were first experienced in the United States, Europe or other leading countries in Asia. Indirectly, this can be associated with Internet connectivity and the release of Internet devices. Contracts between vendors of the Apache Hadoop (Cloudera, Hortonworks and DataStax) and telecommunications providers are very often signed in the headquarters of the multinationals in Europe or the United States. One example of this is the case of Cloudera and Telefônica. Cloudera signed contracts with Telefônica Spain to solve similar data problems in Brazil and four other countries in Latin America. Móvil operates in 15 Latin American countries. TIM used the same big data solution for TIM Italy and TIM Brasil.

In terms of human resources, there is a scarcity of qualified professionals in Brazil. This reflects the international labour market for highly skilled information technologies professionals. It is important to point out that big data professionals are not necessarily based in an office in São Paulo, London or New York. These types of professionals are distributed around the world and often work remotely as freelancers. As a result, the process of recruiting sometimes differs from traditional industries. It does not necessarily follow the pattern of first and second interviews and a probation period. Quite often, these professionals work on a project basis. Some of these professionals are even able to pick out the most interesting or most profitable projects to take part in. The recruiting process often involves a firm making public an internal software bug – an open call. Candidates are invited to send solutions and the candidates that come up with the best solutions are offered a position. Sometimes there is no solution to the problem, but firms want to challenge professionals. The technology is used by many firms across industries and countries, and the best professionals are often committers of a module and well-known members of their communities. Thus, recruiters sometimes try to locate the main committers to an Apache Hadoop module and offer them a job.

In Brazil, one way Cloudera was able to find human resources was by partnering with F5, a domestic big data analytics service provider. Cloudera licensed some of their courses on the Apache Hadoop to F5. F5 began to train professionals. In parallel to the courses offered, Semantics created a consulting subdivision. While individuals were being trained for a fee, Semantics was scouting for the best talent. The best students were later offered a position at F5. F5 became the domestic Cloudera representative and has replicated the concept in other countries in the South American market.

The fact that there are no university courses in big data or analytics in Brazil has a big impact on the availability of human resources. Students and professionals often go overseas or do online courses. These types of courses can be very beneficial for firms and participants. However, it is rather difficult to replicate most of the processes shown in the course. This is down to three reasons: (1) the in-house hardware is not suitable for such operations; (2) there are no human resources to maintain infrastructure; (3) sometimes compliance plays an important role for firms using the Apache Hadoop. Compliance is discussed in the next case story section (financial sector) in more detail. Additionally, firms often train new members of staff internally and look for adequate candidates in other fields. Many employees come from a physics, mathematics or engineering background. What we can observe here is that in addition to the economic and technological gap between Europe and Brazil, there is also a gap in training.

6.2.2. Community Distribution

The Apache Hadoop consists of two distributions – community and enterprise. The community distribution is free, often consists of newly released codes, and needs much editing. The enterprise distribution is made up of matured codes that have been edited over a long period. One of the features that make the community distribution appealing is the conditions under which the codes can be edited. The Apache Hadoop is based on free open-source software technologies. Free and open-source qualities and structure are discussed in the literature review under Section 3.2.3. The Economy of F/OSS: Ties & Networks.

The community distribution has a significant impact on the development of enterprise distribution in F8. The possibility of testing with the community distribution is crucial for the organisations. F8 must be able to go through a test phase where the team can edit codes and tune the architecture for better performance. During the period of the fieldwork in 2016, F8 was going through an intensive testing process. F8 invested in infrastructure and in human resources to acquire the necessary know-how. Most pilot projects in F8 have started this way – frequently using the community distribution. Through testing, the firm could develop a very specific workflow. Consequently, F8 built on the capacity to outflow knowledge and commercialise the Apache Hadoop to other firms. It reflects outbound OI processes discussed in Section 2.1.5. Outbound OI. In terms of maintenance, F8 still relies on Cloudera, the vendor, due to compliance reasons. The importance of compliance will be discussed in more detail in the following case story, 6.4 Financial Sector. Cloudera and other vendors are discussed in Section 5.5. The Economics of the *Apache Hadoop*.

The examples of F8 illustrate that at the beginning of a firm's journey through the deep blue sea of big data, it is crucial to fine-tune codes to integrate the technology into the internal processes of the organisations.

As said by the big data architect: "Big data entered as a pilot project in our firm with a community distribution. We went out shopping for hardware and started with an internal project with our own data.''

The innovation process in F8 began with community distribution due to an internal data problem. It was necessary to go through an integration period where modules were fine-tuned and first put into operation in F8.

As the F8's solutions architect highlights: *"We do not necessarily have an answer for all issues as they occur. But we have a running Apache Hadoop that is getting more mature at this point. We are now entering the processing phase, analysis, and delivering an answer to our customers. We are exploring a new horizon here in Brazil."*

Projects within the Apache Hadoop are maintained by their main committers, usually vendors. There is a committee within the vendors (Cloudera, Hortonworks and DataStax) who are responsible for analysing the community distribution and determining how stable the distribution is. The criteria seem to be unknown, although there are parameters that seem to be relevant, e.g., adoption and bugs. These are very complex processes due to the direct impact on businesses.

The foundation of the enterprise is the community distribution, meaning everything that is developed in the community ends up migrating to the enterprise distribution. On the one hand, the contribution to the enterprise is imperative in terms of patch codes and change. On the other hand, the development of the official community team is crucial for the organic continuation of the ecosystems. It is the trade-off between what is possible technically – the flexibility of the platform – and what makes sense economically – the operability of the platform. F12 is a leading domestic telecommunications provider. Its operation is community distribution based.

As the MVP[16] of F12 explains:

"The community distribution is always a few releases ahead. Committers release new features in the community distribution at a much faster pace, while the enterprise distribution is, preferably, updated continuously. The slow process enables the

16A so-called most valuable player (MVP). MVP is the title given to select individuals who have made a substantial contribution to an open source project.

platform to mature and become stable over time. In a way, the community distribution is the basis for progression of the whole ecosystem."

Community distribution offers flexibility over operability. When new models are put into operation, there is a need for a comprehensive test phase. Within this period of adaptation, firms go through a process of learning. The insight given by F8 points to two specific characteristics of the community distribution: (1) there is an interplay between the flexibility of codes and how firms innovate, and (2) there is a process of code migration from the community to the enterprise distribution.

F12 deploys the Apache Hadoop for a very specific application that is used for high-performance and high availability. The updates in the system's core must improve the stability and availability of the platform. Along these lines, the vendor did a great job of rewriting the heart of one of the Apache Hadoop's modules (Apache Cassandra). The types of improvements add a lot to F12. They reduce the resources needed and do more work with less infrastructure. The improvement is significant in financial terms. The Apache Cassandra is discussed in Section 6.3. What is the Apache Hadoop?

One main advantage of the enterprise is support. Depending on how many nodes the organisation has, the firm can pay for emergency support around the clock. The participant claims that the implementation of the *Apache Hadoop* is not a trivial process, and it may not work out, which happens a lot. Enterprise distribution is a good alternative since firms have the security of having a trained professional oversee the project. Although F12 paid for a licence from 2009 to 2012 (enterprise distribution), it deployed the community version with paid support. F12 was one of the pioneers in using Apache Hadoop in Brazil.

The MVP added:

"We were forced to pay for the licence! To get support and have a specialist helping us 24/7. We used Apache Hadoop for a critical service: pricing. So, it's a service that has to be impeccable. It can never fail; otherwise, the company's revenue falls proportionally. It's a thing that can't crash; it can't slow down."

The example of F12 illustrates that the process of implementation and adoption is very complicated. Firms that have human resources are sometimes inclined to pay for licences to understand the code better. It is vital to take into account that the Apache Hadoop is composed of hundreds of small components. The vendors are very skilled at putting different open-source components together and out-bounding them in a package. The primary competence of vendors is to solve these kinds of problems. F12 shows that it had the human resources and the know-how to handle the Apache Hadoop. However, the codebase offered by the vendor was more operational. It integrates very well with other components of F12's workflows. Once F12 understood the codes, it transferred the new knowledge to its community distribution.

The example of F12 reflects the extensive literature on outbound OI on licensing. Yet, F12 stand out from licensing seen in other industries, e.g., the pharmaceutical industry. Because F12 favoured using the community distribution, it differs slightly from examples seen in firms using an enterprise distribution for operations. As suggested by F12, the licence was acquired to better understand the codes.

The same process was also observed in F8, which needed to use a tool from the community distribution to improve procedures within the firm. It was necessary to buy the licence and negotiate with the vendor to analyse the codebase of the enterprise distribution and eventually upgrade the firm's community distribution. There are some features for which a firm would need a licence but not necessarily if the firm could work out the code by itself.

Within F8, there are no internal forces towards the acquisition of an enterprise distribution, but it is very likely to acquire licences and contracts with vendors to develop a more polished codebase shortly. The better the codebase, the better the performance.

F3 went through similar processes to F8 and F11. It deployed the very first release of the Apache Hadoop – without the support of a vendor. Although the enterprise distribution offers support and easy integration, F3 was looking for knowledge instead. On one particular occasion, it discovered a lot of bugs and problems and reported them almost weekly. This was beneficial as it enabled F3 to gain and exchange knowledge and ideas. Additionally, F3 had a lot of problems as the literature in Portuguese on the Apache Hadoop was almost inexistent. F3 built up the knowledge needed as a team. It had other employees from England who had implemented the Apache Hadoop before and helped the Brazilian subsidiaries with the task.

In less than six months, several other tools emerged. However, F3 could not use them because the tools were not stable enough. Some modules were released very recently in the community distribution but they could only be used while conducting a test. F3 understood that it was always better to wait for a later release. Usually, the older the release the more robust it is. The examples of F3, F7 and F11 show that it is complicated to integrate the newest releases with other modules of the Apache Hadoop. It needs much tuning before it can be used within a running system.

This section has discussed the examples of F8 and F12 that bought licences to look into the codebase of the *Apache Hadoop* but not necessarily for operation. According to F12, it was very keen to use the newest release of an *Apache Hadoop* module (*Apache Cassandra*). The *Apache Cassandra* is discussed in Section 6.3. What is the *Apache Hadoop*? The latest release of the *Apache Cassandra* (*Apache Cassandra 2.0*) was rewritten from scratch. It gained a lot of stability and some commercial features that F12 was very interested in using. DataStax added operational concepts that have dominated the database industry for years – concepts that F12 was very familiar with. By making the commercial feature look and feel like concepts that dominated the data sector for years, the vendor aimed at embracing a more significant market share.

Table 52. Take-aways from the Community Distribution in the Telecommunications Sector

Take-out	Firm	Motivation
The importance of the community distribution for the development of the enterprise distribution	3 8 12	Testing the performance of the *Apache Hadoop* before migrating codes to the enterprise distribution. Most projects within F3 start with community distribution. Outbound the firm's technology and knowledge to other firms. Most pilot projects start with community distribution. The community distribution is essential for the organic continuation of the *Apache Hadoop*.
The community distribution is the foundation of the *Apache Hadoop*	12 8	Committers release new components at a faster pace at the community distribution. The process enables the enterprise distribution to become stable over time. There is a process of code migration from the community distribution to the enterprise distribution.
Reasons to acquire a licence	8 8/12 3	For support at the beginning of the innovation process. Could not work out the code on its own. The firm had to negotiate with the vendor to update the community distribution. The firm did not get a licence because it wanted to acquire knowledge instead.
Outflow of external knowledge of vendors	3/12	There is a committee within vendors that is responsible for determining the stability of the

		distribution.
External pathways to commercialise a technology	12	Due to the implementation of approaches used for decades in databases, the firm's employees had a much smoother transition from the "old" ways of handling data to concepts of big data.
		A team from a European branch helped the Brazilian subsidiary with the integration and implementation of the *Apache Hadoop*.
It can be complicated to integrate a new release to a running *Apache Hadoop*	3 8 12	The firm reported many bugs and corrections during the period of organisational learning. It had to wait for later releases so it could use them in the *Apache Hadoop*. The organisation paid the licence for support only while learning how to integrate the *Apache Hadoop*.

It is also important to note that it was done intentionally as a strategy so that data professionals could make a smoother transition to Apache Cassandra 2.0. It was so successful that DataStax added additional features on the Apache Cassandra 3.0 that would follow the same strategy. Table 52 summarises the take-aways and key applications of the community distribution in the telecommunications sector.

6.2.3. Enterprise Distribution

Section 6.2.1. A Truly International Sector has discussed the scarcity of specialised labour in big data. Brazil reflects the international labour market on qualified human resources and the lack of domestic university degrees aggravates the situation.

This is exemplified by F24. Due to the complexity of the Apache Hadoop and the scarce human resources available in Brazil, F24 decided to implement the enterprise distribution – the community distribution needs a lot of maintenance and support. F24 decided to pay for licences rather than going through the learning curve. According to the firm's calculation, it would cost more to invest in human resources to maintain the platform than for a licence. The Apache Hadoop has to be integrated, and it takes a lot of programming. The firm concluded that this would present numerous difficulties. As the data engineer says:

> *"It was better to pay twenty thousand dollars to make sure that when I have a problem, I'll have someone to call 24 hours a day."*

The use of cloud computing and the partnership with Microsoft Azure follow another very interesting strategy by DataStax to innovate and outbound their knowledge and technology (Apache Cassandra) to other firms. As discussed in the previous example, DataStax used a programming language and a database to attract a target group to its technologies and outflow knowledge and accelerate its technology. In the following example, DataStax will use a strategic partnership with Microsoft to outflow knowledge and commercialise the Apache Cassandra.

DataStax does not offer compliance on cloud services for clients using the *Apache Cassandra*. But Microsoft does – through Microsoft Azure. 95% of the Fortune 500 companies already use the Microsoft cloud computing service. This does not only increase the amount of trust from other firms in the cloud service of Microsoft, but it also gives DataStax a whole new portfolio of clientele. For instance, F8's client was familiar with Microsoft Azure and suffering from scalability. Due to the partnership between DataStax and Microsoft, it was a relatively uncomplicated job for F8 to move the client's data to *Apache*

Cassandra and link it to Microsoft Azure. According to the Solution Architect at F8, this was done with a couple of mouse clicks.

As seen in the previous example, F8 still relies frequently on cloud computing "SaaS" – software as a service – because it is customary for its clients, even for compliance reasons, to do so. There are two lessons to take from the hybrid solution offered by F8: (1) although F8 is capable of consulting and implementing solutions, due to legal reasons, it is not able to sign a service-level agreement with its clients (compliance will be discussed in case study 6.4, Financial Sector); (2) F8 shows that to sell its service (consulting and implementation), it buys licences from vendors so that they can offer a "certain" level of support to clients. F8 is partly responsible for the performance of its clients' set-up and is the first line of contact. Whereas F8 solves minor issues, significant problems are resolved by F8 with the instructions of vendors.

F12 and F8 have demonstrated how vendors enter alliances or introduce technologies to the processes of the *Apache Hadoop* to outflow knowledge or for commercialisation. The next example indicates outbound innovation processes on a smaller scale. F16 is specialised in hardware and the distribution of storage equipment for big data. Over the years, F16 has accumulated a lot of experience in implementing the *Apache Hadoop* and has recognised the opportunity to outflow its knowledge and commercialise the dedicated hardware and big data service to its clients. F16 can install the hardware and push the enterprise distribution on its clients.

As the solution architect highlights:

"I think the enterprise distribution is gaining continuously in market shares. We have diversified and extended our offerings. Today, we sell and install the Apache Hadoop with our hardware all over Brazil."

Table 53 illustrated the take-aways and key applications of Enterprise Distribution in the Telecommunications Sector.

Table 53. Take-aways from Enterprise Distribution in the Telecommunications Sector

Take-out	Firm	Motivation
The community distribution is the foundation of the development of the *Apache Hadoop*.	8	Enterprise distribution contributes to the *Apache Hadoop* in terms of patch codes and change.
Reasons to acquire a licence	24	According to the firm's calculation, it would cost more to invest in human resources and time to go through the learning curve.
Outflow of knowledge or commercialisation of technology	8 12 16 20	The integration of *Apache Cassandra* in the Microsoft cloud service – Microsoft Azure – to overcome the need for compliance of cloud operations. Introduction of established technologies in the workflow of the new release of the *Apache Cassandra* to ease the transition of professionals to big data. The firm outflows its knowledge (enterprise distribution) in combination with the commercialisation of big data hardware. Firms can use cloud services to outflow their knowledge with little resistance from clients.

A similar example was given by F20; it makes no difference to F20 what distribution the clients want. Very recently the demand for cloud computing has increased in F20. The concept of software as a service (SaaS) is very simple to explain to a client. SaaS seems to be an easier way of starting a project within firms that have no track record in big data or *Apache Hadoop*. If a firm wants to invest in big data (*Apache Hadoop*) it requires a lot of commitment to the organisation. It needs financial power as well as dedicated human resources and the ability to use the data available. What F20 has seen over the years is that there are many firms with very useful data who are willing to work with it, but they do not know exactly how to start. Most firms associate big data with a high investment with a long period of no return. The concept of cloud computing, however, is a way to start small and increase investment gradually. As F20's analytics sales executive says:

> "*My clients pay for enterprise distribution because it offers an additional guarantee. It may be security because big data is all about data. It may be scalability or evolution. If the community distribution can't guarantee these key points, my clients end up adhering to the enterprise distribution*."

6.2.4. Conclusion

Data is crucial for organisations in the telecommunications sector, and machine-to-machine communication is about to surpass human-to-human and human-to-machine interaction. Indeed, 95% of organisations in the sector have already adopted big data technologies, and growth is expected with the introduction of 5G technologies. Over the last two decades, the telecommunications sector has witnessed significant technological innovation and regulatory change. Whereas firms in the sector relied on telephony and text messaging, they are now data-centric. Similar to other economies, Brazil has gone through privatisation, which has contributed to the wide diffusion of connectivity in the country.

Multinationals dominate the Brazilian telecommunications market. Therefore, data problems have been experienced first-hand at the headquarters of these firms. Consequently, issues are first solved overseas, and vendors deploy the tested solutions in the Brazilian subsidiaries. Another important take-out from the case study is the lack of qualified big data professionals worldwide. Brazil reflects the same reality. As domestic universities do not offer degrees in data science, it makes it even harder for Brazilian firms to find adequate professionals. Brazilian firms often train their professionals internally, and the recruitment process sometimes differs from that of other industries. One firm was able to use a licensed course by a vendor to hire the best participants.

Some evidence shows the importance of community distribution for the development of enterprise distribution. F3 and F7 have demonstrated that most projects start with community distribution. The firms eventually outflow their knowledge and commercialise the *Apache Hadoop*.

The community distribution is essential for the organic evolution of the *Apache Hadoop*. The telecommunications case study indicates that community distribution is the foundation of the development of the *Apache Hadoop*. Committers release at a faster pace at the community distribution level, and there is a process of code migration from the community distribution to the enterprise distribution. However, the exact parameters seem to be unknown.

The case story also pinpoints different reasons for the acquisition of a licence. F8 needed support at the very beginning of the innovation process. At the same time, F8 and F11 could not work out the codebase and had to negotiate the acquisition of a licence. Whether to acquire a licence or not is due to how vendors explore different pathways to commercialise the *Apache Hadoop*. The example of F11 has shown how the *Apache Cassandra 2.0* has enabled a smoother transition for the firm's professionals. This, in turn, indicates the usage of external pathways to commercialise the *Apache Hadoop* by DataStax. While the community distribution contributes to the organic evolution of the *Apache Hadoop*, the enterprise distribution is critical in terms of patch codes and change. The integration of the *Apache Cassandra* in the cloud service of Microsoft (Microsoft Azure) and the inclusion of established technologies in the *Apache Cassandra* indicate outflows of knowledge and attempts at commercialisation by DataStax.

6.3. FINANCE SECTOR CASE HISTORIES

6.3.1. Community Distribution

The telecommunications case stories discusses the importance of community distribution for the organic development of the *Apache Hadoop*. It also suggests that commercialisation of the *Apache Hadoop* often happens via the enterprise distribution. This case stories has highlighted the importance of compliance for the financial sector.

Financial institutions often look for new features in the community distribution (usually more advanced technically yet unstable). This, in turn, may be a bit difficult to accommodate as the enterprise version is usually a different release. Firms may have to wait and not be able to integrate these features into their services. Such cooperation between vendors and clients may need to be synchronised and negotiated. Here is where the trade-off between flexibility (technically advanced) and operability (stability) plays a crucial role again. On different occasions, F18 has found a module that would be excellent for operations;

however, the module was recently released in the community distribution. As discussed previously, new releases presented some issues and, therefore, were not suitable for integration into F18's running enterprise distribution. F18 communicated to the vendor the need to use the new module, and they worked together on the solution. Eventually, the issues were solved and integrated into the following release of the vendor's enterprise distribution. However, it is not always so straightforward. In some more complex situations, vendors are not able to solve the issues; this is when the communities would be encouraged to solve the problems.

According to F1, there are economic risks if banks lose some bank transactions. Most importantly, the image of a bank can be compromised. Whenever a bank wants some functionality that is still in the community distribution, banks contact the vendors. Vendors add the solution to the subsequent release of the enterprise distribution. Once the feature is added to the enterprise distribution, everything is governed by contracts. If there were financial losses, banks are covered by their insurance (this is why there are no commercial or investment banks, small or large, that use open-source on banking systems platforms). Community distribution has been used for human resources platforms or less important processes. However, where businesses take place, it is improbable.

The previous case study (telecommunications) has discussed examples of firms deploying the community distribution for operations. There were even examples of firms buying the licence only to have a look at the codebase. The telecommunications sector has shown that while the enterprise distribution plays an essential role in the process of the maturation of the codes, firms favour the flexibility of the community distribution for operations.

It does not reflect the circumstances under which firms operate in the financial sector. F19 often uses the community distribution for testing, proof of concepts, and the deployment of new workflows. As a consultancy, F19 works very closely with vendors (Cloudera and Hortonworks) in many different projects for different clients to tune the Apache Hadoop before delivering the final solutions.

As the big data engineer of F19 says:

> *“The Apache Hadoop involves several projects. These projects are developed and maintained by third parties like us. We test and correct issues and often tune the Apache Hadoop for best performance with the community distribution. Later, when the solution for a particular problem has achieved a satisfactory result, we incorporate the solution into the clients' enterprise distributions.”*

The same was observed in F16 and F20. The firms use the community distribution for development, prototyping and testing. For instance, it is much easier for Firm 16 to offer analytics services and consulting as it used to before the Apache Hadoop. The foremost opportunity with the community distribution is the possibility of selling analytics services with a prototype only. The community distribution is the primary catalyst, along with cloud tools. The two technologies have significantly increased and potentiate the selling of analytics projects in F16, particularly analytics services and the creation of new big data systems. F16 deals with very particular situations and problems that require very specific troubleshooting. Once the problem has been solved, he tries to identify other industries that are going through similar data problems and tries to scope for new projects in different segments.

According to Firm 20, big data is a new concept in Brazil. The idea of data for businesses is still in its infancy; therefore, some firms are reluctant to invest in big data technologies. The community distribution enables firms or those working in consulting to develop a project and present the results. Depending on the needs of the clients, it is possible to migrate to an enterprise distribution.

As the analytics sales executive of F20 explains:

"It is tough to justify an investment. There is a big problem in investing in big data analysis platforms. The difficulty lies in assessing the direct impact of the investment and the financial return."

F17 has highlighted the same difficulties regarding the ability and willingness of domestic firms to invest in the *Apache Hadoop*. Most projects in Brazil are still for cost-cutting. There is still a barrier to breaking paradigms because most firms in the sector are not used to programming and the management of high volumes of data. The "old" ways are still dominant. Most organisations are not using big data because the costs of investing in new technologies and human resources are very high. The high costs result in new openings for smaller firms. Many consultancy firms are offering assistance with the very *Apache Hadoop* they have developed over the years, which is almost 100% community distribution based. These firms do not have to make a significant financial investment in tools and infrastructure. It is possible to rent cloud computing for a reasonable price and install a lot of community distribution tools, meaning lower entry-level and smaller firms can compete with many big analytics firms. There are smaller firms and start-ups offering consultancies, implementation but maintenance due to the service-level agreement (SLA). SLA is closely linked to compliance.

The reality is reflected by F9. The firm has specialised in the integration of big data technologies in firms in the financial sector. F9 works with both Apache Hadoop distributions (community and enterprise) and believes the enterprise distribution has the support that serves its clients very aggressively at a high level in terms of batch correction and stability. According to the solution architect of F9, the community distribution is free of this commitment; however, there is what the participant would describe as the trade-off or cost-benefit, as a subscription is extremely expensive. As he says:

> *"Vendors (Cloudera, Hortonworks and DataStax) tie incomprehensible business models to their offers. It does not make sense at all as the cost of managing five nodes will not differ greatly from the cost of managing ten. It explains the vacuum left for new organisations (smaller starts-up) to fill in the gap and compete with the major distributors."*

Table 54 illustrated take-aways from the community distribution in the financial sector in Brazil

Table 54. Take-aways from the Community Distribution in the Financial Sector in Brazil

Take-away	Firm	Motivation	Reason
Importance of syncing the two distributions	25 1	Offering clients solutions to existing problems Preventing fines and risking image damage	Community distribution is always some releases ahead of the enterprise distribution. Communication between vendors and their clients is essential. Banks are only covered by their insurance if they deploy for businesses enterprise distribution with SLA.
Testing and proof of concept	16 19 20	Tuning the *Apache Hadoop* before delivery Selling analytics services with a prototype Breaking a paradigm in firms in the sector in Brazil	It takes time to achieve a satisfactory result with the *Apache Hadoop*, and tests are essential for the best performance. Consultants can show results with lower costs with the commitment of clients. Some firms are reluctant to invest in big data technologies. Most firms are still not using big data because doing things the “old” way is judged as more cost-effective.
New opportunities for start-ups and smaller firms	17 9 25	Filling the gap of established vendors The price and the incomprehensible business models of vendors Smaller projects that do not need 24/7 service	Smaller firms can offer the very *Apache Hadoop* they have developed over the years. No need for service-level agreement (SLA) and compliance. Cloud computing can help smaller firms to compete with big vendors

The sector, however, is governed by compliance and regulations. There are still firms that do not fit under this umbrella and do not operate under such restrictions. This is the case of F25. The organisation is active in social credit scoring and is not constrained to a particular distribution, compliance or vendors. F25 has no demand for enterprise distribution as the project has neither the dimensions nor the need for 24/7 maintenance. F25 does not need a service-level agreement (SLA).

As the firm's consultant explains:

"I don't work with platforms that need SLA. The only thing I need is computational power and networks."

Table 51 summarises the key applications and take-aways of the community distribution in the financial sector in Brazil.

6.3.2. Enterprise Distribution

Cost plays an essential role in firms when deciding what distribution to use. As seen with F9, it can be costly to run an enterprise distribution. Furthermore, the cost associated with it is not always justified; however, not all firms in the analysis share this position.

F23 uses enterprise distribution due to the cost of human resources. The community distribution requires much maintenance and support, and F23 decided to pay for licences rather than going through the learning curve. According to the firm's calculation, it would cost more to invest in human resources to maintain the platform. For the *Apache Hadoop* to function it has to be integrated, and this takes much programming. F23 saw that this would present numerous difficulties. As the firm's software developer says:

"It was better to pay twenty thousand dollars to make sure that when I have a problem, I'll have someone to call 24 hours a day."

Enterprise distribution dominates operations in the sector. Section 6.3.1. International Agreements and Standards introduces legal and international bodies that suggest and enforce some guidelines for organisations. F5 reflects reality.

The company acts in the interests of fraud prevention. Before a transaction is authorised the Apache Hadoop checks if the purchase is legitimate or not in what the firm describes as "classification templates". F5 works with the enterprise version only due to the high level of service-level agreements (SLA). There are millions of transactions a minute, 365 days a year. Each sale is important, and the fines associated with technical failure are high.

The example of F23 has shown that enterprise distribution can help a firm with no human resources to take advantage of big data technologies. F9 justifies the use of enterprise distribution differently. While human resources are available, the lack of partnership with local firms seems to narrow the option of F23. The idea of developing new revenue streams is about offering third parties' software-as-a-service (SaaS). When a firm has the building blocks of the platform, the computing network, and code versioning (enterprise distribution), there is still the need to orchestrate and coordinate the creation of the service. The service can be an application or a platform. In the segment of big data, repeatable scalability is crucial. F9 highlights the difficulty of finding a domestic cloud partner. Therefore, the best option at the moment is to work with the enterprise distribution.

F9's developer concludes:

"I am a developer and I don't understand cloud servers. I just code. I need a partner that sets the cloud parameters and manages it for me."

Section 6.3.3. International Agreements and Standards introduces legal and international bodies that suggest and enforce some guidelines for firms. F1 is one of the firms that is regulated by the agencies discussed. F1's marketing data department manages all the data operations that take place within the bank. It is from there that all information goes out to Brazilian and European regulatory authorities and for risk assessments. The department handles a colossal amount of data.

Consequently, F1 has created an internal department to prevent friction with the regulatory authorities. F1 is obliged by its insurance to work with enterprise distribution. However, if F1 would only consider the financial aspects, it would instead work with the community distribution.

As the data analyst and scrum master said:

"*I think it's much cheaper to work with a community distribution and to have dialogue with the communities to evolve the Apache Hadoop. The community distribution offers many ready-made solutions that are easily scalable. However, it's is very complicated to implement a commercial tool and explain to clients that a service is not working because the bank has used a free solution – community distribution*."

Table 55 summarises take-aways from enterprise distribution in the financial sector

Table 55. Take-aways from Enterprise Distribution in the Financial Sector

Take-aways	Firm	Motivation
Cost benefits of the enterprise distribution	23 5 9 1	The community distribution requires much maintenance and support. The fines associated with technical failures are very high. Lack of domestic cloud computing partners. It is important to prevent issues with authorities and keep the image of the entity intact.

6.3.3. Conclusion

The data analysis has shown that the community is always some releases ahead of the enterprise. The community distribution is very often used for testing and proof of concepts and is rarely deployed for operations due to compliance. The fines associated with technical failures are very high. The community distribution enables consultancy agencies to present projects to clients with no need for significant financial commitment. Firms can suggest solutions developed in-house over the years. Firms such as start-ups or smaller consultancy agencies are filling the gap of established vendors because they are too expensive for smaller firms with limited financial resources. That is why financial institutions not only obey national legislation but are also under strict regulation by international bodies. Such international agencies advise their members on different risk issues. Of interest here are the operational risks associated with IT infrastructures published by the Basel Committee on Banking Supervision in 2003 under the document of BCBS 23923.

6.4. RETAIL AND MARKETING CASE HISTORIES

6.4.1. Community Distribution

The previous case studies have discussed different examples of how firms have used the Apache Hadoop. There was an example of a firm that wanted to go through the learning curve. Other examples include firms that wish to conduct tests before migrating to the enterprise distribution, and firms which use the Apache Hadoop for operations.

F5 is a domestic online classified advertisement website – the equivalent of Gumtree in Britain. In 2016, F5 was relatively new to Apache Hadoop and big data analysis. During the period of the fieldwork, F5 was trying to understand how to fund and develop an Apache Hadoop project. The example of F5 indicates that it is not only about how to use Apache Hadoop in technical terms. First and foremost, firms need to understand how to finance the integration of big data technologies in the firm's workflows and service portfolio.

The last two case studies in this chapter have covered many different difficulties related to the operations of the community distribution. F11 has used the enterprise distribution at the beginning of the innovation circle; eventually, F11 put the community distribution into operation. F3 has decided to invest time and went through the learning curve. Other examples have shown that firms use the community distribution for testing or to show results with little financial commitment from the clients. But, what makes the example of F5 stand out from others seen so far in this chapter is that F5 was confronted with technical and financial 'adaptation'. What can be interpreted here is that firms often have limited resources, and these have to be reorganised. According to F5, it is sometimes easier to invest in enterprise distribution with support. Vendors can bring solutions to specific problems in a shorter period. In the particular case of F5, the firm did not only have to relocate human resources to master the Apache Hadoop technically, but it was also confronted with financial questions. Do we need more labour? Will we have to pay for support? How long are we going to test for? Finally, how much will it cost us?

As opposed to F5, F21 is an IT multinational consultancy with a subsidiary in São Paulo. Unlike other firms, it has no problem financing an open-source IT project because open source is deep-rooted in F21's philosophy. F21 is an organisation that has grown together with the open-source software movement. Today, F21 offers consultancy in cybersecurity, digital transformation and big data services for industry, wholesale and retail.

F21 is often confronted with clients that are unaware of the financial benefits big data technologies can bring to an organisation. In F21's view, this has a lot do to with how organisations are used to predicting the growth and returns of their investments. For instance, if a firm hires new people to improve a product or service, the firm will potentially have a 20% increase in profitability. This will depend on various external and internal factors which will be not discussed in depth here.

With big data, the logic is different. Big data is about recommendation models, algorithms and predictions. For example, if a firm makes a model and the model is successful, potentially a firm can improve sales by 30%. Brazilian firms are often chained to the 'old' logic of organisational thinking. So, it is challenging to justify and assess the impact of big data investment. It is a new way of thinking for many firms in Brazil. At the moment, many firms in Brazil are afraid to make investments, no matter how small. The exciting thing is that the community distribution can help firms overcome fear. F21 can develop projects, present the results and from there, depending on the needs of the client, F21 can suggest something more appropriate to the client's need.

What makes the examples F5 and F21 very compelling for this analysis are the financial aspects that are often taken for granted. This chapter has discussed the lack of human resources, how the technology was imported from international headquarters, compliance, and the difficulties in implementing new releases in a running system. F21's clients are firms which are new to big data and the Apache Hadoop. They have embedded concepts of what the return of investments should look like in order to plan to create growth strategies. Most importantly, they ask how the high investment can be justified in front of the director's board or CEOs and investors. The example of F5 illustrates an organisation with limited resources and a lack of technical skills. F5 is an online retailer that wants to use big data tools to improve organisational performance to be more competitive. Table 56 summarises take-ways of community distribution in retail & marketing sector.

Table 56. Take-ways of Community Distribution in Retail & Marketing Sectors

Take-aways	Firm	Motivation
The importance of community distribution for firms to learn how to fund an Apache Hadoop project with a lack of technical skills and limited financial strength.	5	The reorganisation of resources within firms can be as challenging as the technical issues. The firm had limited resources and was confronted with questions that bigger firms take for granted sometimes.
Resistance to change organisational thinking	21	It is challenging to justify the high investment in the Apache Hadoop. This has a lot to do with embedded organisational thinking of how the prediction of returns and growth is carried out.
Opportunities for start-ups due to high commissions The process of innovation in starts-up can take different forms, as seen in bigger firms	23	Due to the background of the founders, the organisation was capable of overcoming any challenging financial aspects. This highlights the low financial entry requirements of the Apache Hadoop. The size of the firm combined with the flexibility of community distribution enabled the firm to explore different ways to use the tools at hand.

This lack of resources and technical skill is not the case of F23. F23's co-founder and CEO is a big data analyst with a BA in computer science from a foreign university. F23 is a start-up specialising in developing online shopping platforms. The start-up combines big data with artificial intelligence and uses algorithms to increase engagements rates.

As F23's CEO says:

"The process of creating cost-effective solutions using community distribution is a process that was made by people who can solve problems using a technology that was developed to solve a different problem."

This analysis has discussed that vendors (Cloudera, Hortonworks and DataStax) offer excellent service and support (SLA), but these can be very cost-intensive. Although some firms cannot afford to buy a license, the Apache Hadoop is very beneficial for their operations, which has resulted in a gap that has been filled by starts-up like F23 or other organisations that have developed their own community distribution over the year. It is important to note that F23's main business model is not the Apache Hadoop itself but data analytics. F23's biggest challenge is to transform the technologies the firm has developed into business opportunities.

What makes the example of F23 compelling is the flexibility that community distribution offers the organisation. In this particular example, F23 does not follow the patterns seen previously, which is tuning codes to better performance. The tuning of codes is a continuous process in most organisations using the Apache Hadoop, including F23. F23's CEO highlighted that F23 can use community distribution and try to solve problems with a technology that was developed to address other problems. According to the CEO, problems include the lack of communication between departments within firms, in particular multinationals. As the CEO adds:

"*We need more people with business vision. Technology is a means to an end, not the end itself. I'm not sure if eventually we will need to bring more technology to the* *curriculum of a business person, or bring more business to the curriculum of technology person. But we need to bridge the two professions."*

6.4.2. Enterprise Distribution

Section 6.2.3. A Truly International Sector has discussed how data problems in telecommunication service providers are first solved at headquarters in the United States and Europe. Eventually, the solutions are implemented. According to F5, the domestic market is not yet fully developed, or, at least, it is not big enough. Big data solutions are mostly seen in large organisations – banking or telecommunications. Often, these firms have in-house solutions with the support of one of the major vendors (Cloudera, Hortonworks and DataStax). The vendors are discussed in Section 5.6. - Apchqa Hadoop Vendors.

Vendors target their sales efforts to bigger firms because if they tried to offer solutions for smaller firms, they would have to move to more specific answers. When a small firm decides to pay for enterprise distribution, it often pays for many tools and features that the firm does not need. This was the difficulty that F5 experienced in 2016, simply because enterprise distribution is not motivated by solving specific and small business problems but rather the big ones. F5 has a small budget, and its major problem was that it was looking at solving particular data problems while the enterprise distribution is better suited to a large-scale solution.

F5 was quite keen on the IBM Apache Hadoop cloud solution but it is prohibitively expensive, and F5 found it unaffordable. The other issue with the IBM solution is that it is not transparent. It often drives F5 to connect with other IBM solutions. This solution is simply unrealistic in terms of cost for small domestic businesses like F5. As F5's' software engineer and architect says:

"*I think what defines the moment of big data in Brazil is learning. I see a little fear among companies and the Brazilian community to adopt the Apache Hadoop, mainly because of the initial cost and the lack of skilled labour to deploy and manage the Apache Hadoop. There is little space for simple solutions for small businesses like us in the strategy of vendors.*"

According to the F15 website, the organisation is the biggest online classified advertisement website in Brazil. The organisation has over 700 employees working on how to bring the right article to the right buyers. While F5 can be compared with Gumtree, F15 could be seen as the Brazilian equivalent to eBay. F5 and F15 use the Apache in the same way. The key application of firms in the sector was discussed in section 6.8.4.3. The Technology. F5 and F15 share very different realities.

As discussed, F5 had difficulties in financing the Apache Hadoop. The firm would favour a more simplified package of enterprise distribution. F15 is financially stable and saw it instead as a technological challenge, not necessarily a financial one. Thus, F15 sees much difference between the two (community and enterprise). According to F15, it is possible to find all tools in both distributions. As F15's data engineer says:

"*We are an organisation that does not sell the Apache Hadoop, but rather consumes the technology. I believe that all features of the Apache Hadoop we use here are in both distributions. If I had to migrate from one to the other, I could do it perfectly, with no major issues. We don't pay a license: we for pay for a service.*"

According to F21, many organisations in Brazil are not very reliable. They make promises to clients that they cannot fulfil. There are consultancy agencies selling services and the concept of big data to other firms but with no skills or technical tools to deliver what they are promising, resulting in serious firms like F21 being discredited in the market. There have been occasions where F21 has had to tell its clients that what they had been offered from its 'competitors' would not work. Unfortunately, firms have said that they can use social media data from Facebook for a marketing campaign for free.

As F21's analytic sales executive says:

"*There is one thing that has to be looked at very carefully here in Brazil. There is a major gap between what customers are being promised and offered, and what they will get and where they will get it from.*"

According to F21's analytic sales executives, big data is a means to an end, and it is not the silver bullet. Many companies new to significant data processes are expecting more results than they can get. The major problem here is that for some operations, firms will have to have the whole infrastructure and data in place. If the two resources are not there, there is no point in investing a lot of money and time. That is why sales executives have to be straightforward from the beginning of the project. F21's analytic sales executive continues:

"*I can contact a company and say that it can be done and let the client know that I may be able to increase their profitability. At the same time, I need to make it clear to them that I depend on several factors – data and tools.*"

F25 is a certified multinational SAP[17] partner and is specialised for business solutions for SMEs. F25 operates subsidiaries across Latin America, the Caribbean, the United States and Europe. The Brazilian subsidiaries consist of a team of ten developers who build mobile applications. As the volume of F25's clients grew, F25 was repeatedly confronted with scalability issues. F25 tried many big data solutions available on the market and decided on

17 SAP (System Application & Products in Data Processing) is a customer relationship management software (CRM) to manage business operations and customer relations.

the Apache Hadoop. F25 judged the Apache Hadoop to be the best solution because the team of developers could build upon already exiting solutions instead of creating everything from scratch. It confirms how free and open-source software development works and highlights why firms are so keen to use (and reuse) the Apache Hadoop. The topic is discussed in Section 3.2.4. Developer's Motivation.

Similar to other firms discussed in the chapter, F25 has used the Apache Hadoop for testing and pilot projects. It turned out that community distribution was too complex to implement, and it needs intensive maintenance and support. F25 has looked into the enterprise model and ended up deciding on it. The reason F25 favoured enterprise distribution was mainly because of the difficulties related to the combination of different modules for different clients. As discussed previously in this chapter, the Apache Hadoop is a combination of more than 100 small components. Taking into consideration that most clients will need a different solution, it could get very complicated for F25. According to F25, they made a commercial decision to avoid problems, prevent risks and limit costs in the future. As the data engineer says:

> *"The problem is that the Apache Hadoop on its own does not answer to all clients' demands. We need the whole ecosystem. It has to be integrated, and it takes a lot of programming and experimenting with scripts and so on. We saw that it would cause a lot of problems with maintenance and production."*

Section 6.2.1. A Truly International Sector has discussed how telecommunication firms have imported solutions from the headquarters into the Brazilian subsidiaries. F8 flew some representatives from a European branch to help the Brazilian subsidiaries to solve a data problem. F25 went through a process of collaboration within branches. According to F25, it would take the Brazilian subsidiaries months to replicate a solution. F25 flew a representative in from the Italian office and solved the problem within two days. As the data engineer adds:

> "*If we had to do it from scratch, it could take up to six months. He did it in two days because he had everything ready! The step by step by step workflow! He knew how to install the environment, how Key Applications of Big Data in the Public Sector to make a change, and to deploy it again.*"

Table 57 illustrates take-aways and key applications of enterprise distribution in retail & marketing sectors.

Table 57. Take-aways and Key Applications of Enterprise Distribution in Retail & Marketing Sectors

Take-aways	Firms	Motivation
Enterprise distribution is mostly seen in big corporations	5	The enterprise distribution covers a wide range of solutions. It makes it hard for a firm looking for concrete solutions to afford it. Firms have to pay for a package that includes a wide range of features that they do not need.
It is easy to migrate features from one distribution to the other	15	Most features can be found in both distributions.
Clients are sometimes promised more than it is possible to deliver for the price	21	There are firms taking advantage of the "buzz" of big data. They promise their clients heaven on earth but cannot deliver.
The community distribution is too labour intensive	25	The firm was not able to handle the integration and maintenance of clients' enterprise distribution. It is far more efficient to buy licenses.

6.4.3. Conclusion

Brazil is one of the biggest retail markets in Latin America, and it has one of the largest marketing sectors in the world. Big data is used in all phases of retail operations and has significantly changed the way marketing campaigns are developed and put into action. The three case stories have shown some similarities regarding the way the Apache Hadoop functions across industries. Firms can understand their clients better; they can improve performance by analysing their operations, and they can improve decision-making.

In various parts of the chapter, there is discussion about the technical difficulties associated with the implementation of the Apache Hadoop. This case study has considered how community distribution can help firms to explore funding options to implement the Apache Hadoop and how resources can be shifted within firms.

There is some resistance to change due to the high investments associated with the Apache Hadoop. Additionally, it is sometimes tough to justify investments. This has to do with embedded organisational thinking about how returns to investments are predicted; it is not only about technological change. Innovation firms, therefore, have to change their organisational thinking.

The case study has highlighted the low entry-level requirements if the firm's owners have a background in computer science. Start-ups are capable of developing their own Apache Hadoop and of exploring different ways to use it. All tools and features are available in both solutions. They are relatively easy to migrate from one distribution to the other. Firms are offering more than they can deliver, resulting in serious firms being discredited in the market.

6.5. PUBLIC SECTOR AND STEEL INDUSTRY CASE HISTORIES

6.5.1. Community Distribution

F2 is a multinational specialised in hardware and has become a Hadoop service owner over the years. F2 works with clients from across industries, including the sectors of aviation, security, metallurgy and health. As service offering has become increasingly relevant for the organisation, it has divided operations into two departments, e.g., hardware and software. The head of analytics and data manager belongs to the software department – what internally is referred to as service. He is responsible for clients' transition from what he would describe as the 'old world' to big data. His primary responsibility is to rebuild and adapt clients' infrastructures to the demands of big data analytics.

F2 has advanced very different applications of the Apache Hadoop and has expanded its service portfolio through internal development and acquisition of start-ups and SMEs. One of the most compelling cases is a video analytics platform that is used to prevent terror attacks or small crimes in underground and train stations. The system includes face recognition software and checks for suspected individuals in dedicated databases. It is a similar system to the one seen at Tottenham Court Road tube station or Heathrow Airport in London. The system analyses if a passenger acts suspiciously while walking along the platform or if the person stands on a platform for a long time without catching a tube or train. It can even check the temperature of passengers to determine if they are nervous before walking through customs.

F2 is very compelling for analysis because it shows that multinationals are in direct competition with vendors (Cloudera, Hortonworks and DataStax). The main difference between the traditional vendors and multinationals offering the Apache Hadoop is that in most cases, multinationals do not provide their Apache Hadoop for download over the internet; instead, they outbound the technology though software as a service platform. Vendors are discussed in detail in section 5.4. The Economics of the Apache Hadoop.

The telecommunication case study has discussed the example of DataStax going into a partnership with Microsoft to commercialise and outbound the Apache Cassandra. It was very beneficial for DataStax to join Microsoft in the project and vice-versa. On the one hand, DataStax could profit from the knowledge of Microsoft in terms of cloud computing, compliance and Microsoft's portfolio of clients. On the other hand, Microsoft could benefit from the distributed server technology and knowledge, as well as the codebase of the Apache Cassandra. It was a win-win situation for both parties.

F2 can very well be compared with Microsoft in many aspects – size, scope and portfolio of clients. But, instead of going into partnership with a vendor, F2 decided to develop its own Hadoop as a service. For someone not immersed in the Apache Hadoop, cloud computing and open-source, it makes little difference as long as it is an Apache Hadoop. However, they are three different things, and this makes a huge difference for organisations managing their outbound OI processes. The fundamental freedoms of free and open-source software play a central role in understanding the two different strategies. Section 2.3. Open Sources discusses the four fundamentals[18] freedoms of free and open-source software, as stated by the Free Software Foundation (FSF).

There is an important reason why F2 has opted to commercialise the Apache Hadoop as a service instead of licensing their technologies and this will be clear after the following insights given by F2's head of analytics and data manager. According to the participant, a few years before the fieldwork, one of F2's important clients wanted to implement the Apache Hadoop. Because the client had previously worked with F2 on other projects over the years, the client wanted to have F2 on-board the Apache Hadoop project too. The client knew that F2 was involved in the Hadoop as a service. The problem was that the client wanted to have the Apache Hadoop installed, along with support, with the service-level agreement of F2. The head of analytics and data manager ended up suggesting his bosses take on the project, but it did not happen.

18 The four fundamental freedoms are: (1) Freedom to use the software for any purpose. (2) Access to the source code. (3) Freedom to make copies and redistribute them. (4) Freedom to distribute the modified version to others.

In order for F2 to be able to offer a service-level agreement, they have to be 100% in control of the codes. Most importantly, because the Apache Hadoop is open-source software, everyone has to have access to the source code and be able to modify it – as stated by the Free Software Foundation (FSF). F2 can overcome the fundamental rights of free and open-source software by offering the Hadoop as a service. As the Head of Analytics & Data Manager says:

> "*Over the years, we have been repeatedly asked by our clients for support and SLA, but we can't offer these because we don't have the level of governance of the Apache Hadoop required. Our firm follows strict compliance rules. The only way we could offer that type of service would be by taking 100% control of the source code.*"

What is important to take from the insights given by F2's head of analytics and data manager is that vendors are 100% in control of the source of the Apache Hadoop. The level of understanding of the source by vendors differs greatly from other firms (multinationals). The example of Hadoop as a service used by many multinationals (Microsoft, F2, Hitachi, Amazon and IBM) highlights once more the importance of vendors (Cloudera, Hortonworks and DataStax) for the advancement of community distribution. Multinationals such as F2 use the recommended version of community distribution for their Hadoop as a service.

Section 6.2.1. A Truly International Sector discussed human resources issues organisations have to overcome. The section has also discussed how many Brazilians are educated in big data overseas. F26 confirms a model used relatively often in outbound OI processes by firms – spin-outs. The example of F26 is not a 'spin-out' in the classical sense seen in the outbound OI literature. The literature discusses how start-ups spin-out because the service or product do not match the main business models of the firm. F26 exemplifies employees of a large firm joining forces and becoming independent. F26 consists of young entrepreneurs who met at F19.

F26 is a start-up primarily specialised in analytics for the financial sector. However, the most compelling user case of F26 is in the sector of construction. The retail and marketing case story has discussed how F23 (due to its size) can use the ‘flexibility’ of community distribution to explore new ways to use the Apache Hadoop to solve old problems. F26 used open data (city hall registration of new properties) in combination with social media (offers of new houses and apartments in construction) to help large-scale manufacture of construction products. The organisation was finding it difficult to predict where a new construction project would be launched. Depending on the location of the project and the organisation managing it, the specification of the properties and the materials used are made up of different qualities or components. The manufacturer of construction products had an advantage over its competitors because it could preview what quality and components to produce and stock its warehouses. The manufacturer was even in the position of calling its clients to make a business offer.

The financial sector has discussed how Brazil has adhered to the European Union recommendation on data privacy released under the General Data Protection Regulation (GDPR). F6 is a multinational focused on the distribution of free and open source technology to the public sector in Brazil and works uniquely with community distribution. During the fieldwork, the Apache Hadoop was being implemented in different departments of the Brazilian government, and F6 was developing a model to serve the domestic public market. Due to Brazilian legislation and bureaucracy, it is a very complex environment. According to F6, it lacks transparency. There were forces within the government that wanted to use different technologies to address the issue.

Section 6.2.1. A Truly International Sector discussed how technologies are developed overseas and eventually implemented by multinationals in Brazil. In the context of the analysis, we considered how telecommunication giants answered to their data problems in the headquarters, and the subsidiaries in Brazil used the same solutions. According to F6, the Brazilian public sector is reactive, in the sense of reusing technologies from overseas. The public sector in Brazil lies far behind the private sector 'technologically'. Accordingly, long negotiation periods are required for what is pretty much routine in the private sector. One of the issues that F6 is continuously confronted with is that the Brazilian government is only interested in pre-tested solutions and does not want to invest in research and training. The private sector ends up adopting innovation from overseas a little faster. During the period of fieldwork, the Brazilian government was still overwhelmed by the benefits of big data and the Apache Hadoop. As F6's Senior Sales Executive explains:

> *"After the major progress that has been made in the private sector over the last couple of years, I think the Brazilian government has acknowledged the potential benefits of big data. In fact, this interest does not only apply to the public sector: there is a greater interest in the Apache Hadoop in general. The technology is gaining relevance across industries in Brazil."*

The insights above point to certain dominance of the community distribution in the domestic public sector. This is mainly due to two factors: (1) it may belong to the culture of the Brazilian government since the introduction of free and open-source technologies by the government of Lula just over two decades ago. (2) The example of F6 has emphasised that the sector is conservative and very seldom applies 'new' technologies. Using the recommended community distribution, specifically the one used by the enterprise distribution, is very cost-effective. The analysis suggests that the two distributions share the same code base. As the code base of the community distribution becomes more stable over the years, it becomes more operational. Although the government handles sensitive data, it does not necessarily need SLA.

The community distribution is used across the Brazilian public sector, not only at federal government level but also at state level. According to an academic at F11, the state of São Paulo is also community-based when it comes to big data and the Apache Hadoop. F11 is one of the most relevant professional education institutions in the country, and the academic acted as a consultant in the project. The task consisted of moving the data of legal cases older than two years to the Apache Hadoop. Even though the example of F11 is not necessarily very technically advanced, it indicates that all levels of the Brazilian Federation are using the community distribution.

While the use of case F11 is not particularly compelling, the case below is far more interesting. F9 is a well-known South American airline: it was founded in Chile at the beginning of the last century and was privatised after the military dictatorship saw its end in the 1980s. The airline was born out of a joint operation with a Brazilian international airline. F9 is a member of One World Group and flies to over 17 domestic and 129 international destinations.

During the fieldwork, the Apache Hadoop was recently implemented in the organisation and still in its infancy. One of the issues F9 is often confronted with is that the difference between releases is immense, and there is one new release every three months. It is a challenge for F9 to keep track of the new releases. The Apache Hadoop requires a lot of maintenance as the members of the Hadoop team within F9 have to keep fine-tuning the platform. This maintenance makes the Apache Hadoop very difficult to operate because it is very labour intensive. Earlier on, this chapter discussed how qualified labour is scarce in the

Brazilian market. F9 is one of the organisations that have to handle this issue, like many others in the analysis, daily. As the technical account manager explains:

> "*It's very complicated because we're constantly tuning the platform. Sometimes, we carry out a Google search query and have to wait for days for the right feedback. It takes an army of warriors to do the job. But it is a lot of fun because we know that there are not many people out there doing the same thing. It's something cutting edge; this part is cool.*"

Considering the scope of an international airline, one would expect that an organisation in the trade would use the Apache Hadoop in very similar ways to the other case studies that have been covered. The chapter has so far dealt with how organisations can improve performance by analysing their own and their client's data. Other firms in the analysis have used the Apache Hadoop to preview interruptions in their service and analyse investments risks. F9 does this too.

The chapter has demonstrated how there are some similarities in how firms use the Apache. The following use case is compelling because it is not an example someone would expect from a business in the trade. F9 has worked together with local and international health authorities in conjunction with partners in the telecommunication sector in the outbreak of contagious diseases. On a domestic level, flight passenger data are regularly used in combination with social media and open data in the spread of Zika, malaria and dengue fever. On an international level, the template was used in an outbreak of Ebola in 2014 in West Africa.

The user case is particularly compelling for analysis because the Apache Hadoop can be used in many different ways. Unstructured and inter-organisational data have increasingly been used to solve every day problems. It is this aspect that makes the Apache Hadoop so interesting for the analysis of innovation.

The chapter has so far discussed different user stories in a variety of industries. Big data and the Apache Hadoop have been discussed in the chapter, but it is not possible to speak about this technology without at least once mentioning algorithms. In the broadest sense, algorithms are rules of calculations. Considering the Apache Hadoop is open-source technology, algorithms create a competitive advantage for some organisations. F16 is an organisation that has focused its business model primarily on the development of its algorithms. F16 is one of the pioneers of artificial intelligence[19] and machine learning [20]in Brazil for the public sector.

During the fieldwork, F16 was specialised in the public sector. However, F16's CEO was expanding businesses to other sectors of the industry, e.g., the health sector and the sectors of manufacturing and finance. The interesting aspect about F16 is that the organisation is not located in the financial and cultural hubs of Brazil (São Paulo and Rio de Janeiro) or the political capital (Brasília). It is located in the quiet little capital of the southern state of Rio Grande do Sul, Porto Alegre. Rio Grande do Sul is well-known for livestock farming and for the European immigrants that came to Brazil over a century ago. Although the state has one of the highest living standards in the country, it is far away from the technological metropolis in the south-east of the country.

Besides its unusual location for a high-tech firm, there is another interesting aspect of F16. A prevalent issue F16 faces is the lack of understanding by its clients as to why the

19 According to Encyclopedia Britannica (2020) artificial intelligence is: "the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings."
20 According to Encyclopedia Britannica (2020) machine learning is:" in artificial intelligence (a subject within computer science), the discipline concerned with the implementation of computer software that can learn autonomously."

organisation cannot reveal its algorithms. F16's CEO explained that some clients have an issue and are reluctant to close the deal when they have no access (transparency) to the algorithms F16 spent years developing. The CEO says:

> "*Clients don't want to accept the analysis. Sometimes, data analysts are more concerned with data transparency, than with the results themselves!*"

The CEO continued by saying that it would depend on the person with whom the negotiation would take place. If a manager looked at the problem, he would accept the results as long as the data was beneficial for the organisation. But not the data analyst! Analysts want to understand how the information was extracted out of the data. Transparency has gained increasing importance for clients. The CEO continuous:

> "*Transparency is a great concern for our organisation. It is a bigger concern than winning customers. Most importantly, transparency will enable us to keep our current customers. In the end, what matters is sending invoices.*"

Table 58 illustrates the Take-aways and key applications of the community distribution in the public sector and steel industry.

Table 58. Take-aways and Key Applications of the Community Distribution in the Public Sector and Steel Industry.

Take-ways	Firm	Motivation
Acquisition of new technologies		Multinationals expand their service through internal development and acquisition of start-ups.
Multinationals are in direct competition with vendors		Multinationals outbound their technologies and commercialise the Apache Hadoop 'as a service'.
Outbound OI processes of multinationals differ from vendors	2	Multinationals can evade SLA with 'as a service' outbound OI processes.
There are different levels of governance of the Apache Hadoop source code		Multinationals offer Hadoop as a service because they are not 100% in control of the source code whereas vendors outbound the technology and commercialise the Apache Hadoop with community and enterprise distribution models.
Start-ups are more flexible to experiment with the Apache Hadoop	26	Start-ups can use freely available resources (community distribution and open data) to develop services to solve real business problems.
The public sector in Brazil lags far behind 'technologically'	6	Brazilian bureaucracy makes simple negotiation very difficult.
The public sector in Brazil does not have an innovative culture	6/11	he Brazilian government only wants to use already tested solutions. It is very conservative and not willing to try new solutions. However, the Apache Hadoop is increasingly being used at all level of government bodies - federal and state.
The community distribution is labour intensive	9	Firms have to continuously fine-tune the Apache Hadoop for better performance.
Private and public sectors are collaborating to address public health issues.		Unstructured data in combination with inter-organisational data are increasingly being used to address public health or other global issues.
Transparency is key in analytics	16	Clients want to understand how a result has been achieved. The

	rules of calculation (algorithms) are as important as the results.
The Apache Hadoop on its own cannot do much	The Apache is only as efficient as the data and the rules of calculation available to firms.
The community distribution is imperative to cost-effective operation	The Apache Hadoop offers an important tool for firms, but start-ups gain a competitive advantage through the rules of calculation (algorithms).

F16 has a great capacity to innovate through the development of algorithms that offer solutions to problems that did not exist in the market. According to the CEO, what keeps F16 motivated today is improving existing technologies or creating technologies from scratch if necessary. The community distribution has become the fundamental factor for F16 to be able to operate cost-effectively. The importance of F16 for the analysis lies in the fundamental idea and structure of how big data analysis works. One thing, is to have the Apache Hadoop (community & enterprise distribution). But, if the software engineer and the data analysts are not working together and in tune with other departments of the organisation, there are limited changes an organisation can make in order to be successful in the trade. In the four case studies, firms and participants have always highlighted the importance of tuning the Apache Hadoop, but algorithms have been neglected. It took a CEO of an artificial intelligence and machine learning start-up to highlight the value of algorithms and data analysts for an organisation handling big data.

6.5.2. Enterprise Distribution

The telecommunication and financial case stories have considered how firms use the Apache Hadoop for a type of laboratory. The case stories have shown how firms have deployed the Apache Hadoop for a pilot project or to fine-tune the Apache Hadoop before migrating to the community distribution. F2 saw it as an opportunity instead to outbound the technology and outbound the Apache Hadoop. F2 offers so-called 'Hadoop Labs' to its clients. For a monthly fee, F2 offers a full operative Apache Hadoop. It consists of all possible hardware, modules and features. The Hadoop Labs came to fruition after F2 closed a deal with one of the biggest fruit juice firms in Brazil. F2's clients are involved in phases of the supply chain of fruit juice – from agriculture to distribution and sales. The client was not willing to close the deal before they could be sure that all phases of the supply chain were in sync with the Apache Hadoop.

What makes F2's Hadoop Labs stand out from the other 'laboratory' discussed in the chapter is that while F2 was gaining know-how in a new sector (agriculture), F2 was outbounding the technology (in the new industry) but, at the same time, F2 was commercialising the Apache Hadoop. According to F2, almost all firms of a similar magnitude operate this way. If the client wants to work on a project, they pay for the Hadoop Lab, and the next day everything is ready to go F2's head of analytics and data manager sees a big discrepancy between firms in the telecommunication, financial and health sectors and firms in the agriculture and manufacturing industries. According to him, the former three sectors are very reluctant to use cloud services due to the responsibility they carry over the data generated by their clients. The latter two are more familiar to cloud computing. As the head of analytics and data manager says:

> *"There is a gap between SLA and governance! And, we managed to make a business opportunity out of it."*

According to the F2's head of analytics and data manager, during the fieldwork, Brazil was going through hard times economically. This, in turn, makes the Apache Hadoop very attractive for firms because it can improve processes and lower risks. Otherwise, it would be challenging to sell the Apache Hadoop just as a pure technological innovation. For instance, F2 has a long-standing partnership with a client in the steel industry. F2's clients had losses in steel revenue of over 10 % due to international steel quality standards tests. In other words, 10% of the total amount of steel produced by F2's client would not pass quality tests – resulting in financial losses for the client. The steel producer went through several trials and had to make a lot of adjustments at the production plant, with no satisfactory results. On top of the 10% loss of steel, F2's clients ended up losing the time and money invested in improving their performance. The steel-making industry in Brazil is a highly automated process, as are most of the manufacturing processes, but F2's client had no data on the operations.

Within F2, there is a team responsible for storage management and automation (SMA Team). The Head of Analytics & Data Manager's team worked together with the SMA Team on the project. First, the most important thing was to collect data because the data were non-existent. In order to do this, the SMA Team installed over 2000 sensors to measure the temperature and pressure of the steel furnace. Although there was no existing data after the installation, there was data coming in real-time. However, as seen by F16, the data on its own does not do much. F2's client was involved in the project in the second phase where F2's client's steel engineers brought in specific knowledge about steel production. According to the Head of Analytics & Data Manager, the second phase is one of the most important phases because the engineers from the field can answer questions such as: "How does your process work?" "What type of steel are you producing?" "What are the losses on the steel and why?"; "What problem do you think you have?" "In which phase of the production is the problem taking place?" It took some time to improve production because not all data was being captured. Eventually, F2 managed to capture data over six months and brought in data scientists specialised in heavy industry to look at the problem. It was a pure data-centric problem which involved a highly technical solution. The solution was finalised by members of F2's headquarter team from the US. It was a very complex operation.

To summarise, F2 installed sensors, collected data for six months, and brought in data scientists with very specialised knowledge. And now comes the exciting part. Once the whole scene was set, F2 had to put all variables together in an algorithm and let the Apache Hadoop do its job. The Apache Hadoop could simulate the steel production and calculate a model of success probability. Now, instead of losing time and money with hundreds of tests to improve the steel production, F2's client was working together with mathematicians to enhance the algorithm to improve the output of the steel. To cut a long story short, the steel producer could improve steel output and save millions of dollars.

According to F2's head of analytics and data manager, where innovation can take place with big data analytics and the Apache Hadoop is in the sectors of manufacture and agriculture. He explains that Brazil's biggest competitor, and yet its biggest ally, is China. Brazil is under a lot of domestic regulations which China is not. The participant was keen to mention two upcoming projects: one in agriculture and one in mining. Unfortunately, he could not go into further detail because the deal was not closed yet. He summarised:

> *"Analytics in industry and agriculture is where we have to focus because banks and telecom have more or less solved their problems. I see great potential in the two sectors, particularly in agriculture, because our industry suffers many tax burdens."*

Table 59 summarises the take-aways and key applications of the enterprise distribution in manufacturing sectors.

Table 59. Take-outs and Key Applications of the Enterprise Distribution in the Public Sector Steel Industry.

Take-outs	Firm	Motivation
Hadoop Labs as a source to outbound the Apache Hadoop Agriculture and manufacture are more accustomed to cloud computing Brazilian firms are interested in results, not innovation Innovation is more likely to take place in manufacture and agriculture Multinationals outbound the technology and commercialise the Apache Hadoop in large scale projects	2	Whereas a consultancy sees it as advantageous to develop a solution with no costs for clients, multinationals charge their clients a monthly fee to use the Hadoop Lab. Telecommunication and financial sectors handle more sensitive data and are reluctant to use cloud-based Apache Hadoop. In the domestic economic situation, firms are willing to invest in technology if they can save money. Most problems have already been solved in the telecommunication and financial sectors. Multinationals are more likely to support large scale projects than vendors.
Documentation needs to be comprehensive and detailed	13	Firms can use documentation in training and for Apache Hadoop configuration.

What can be taken form F2 is that multinationals with the scope of F2 are maybe not able to offer their clients SLA or the options between community and enterprise distribution. Clients are sometimes left with no alternative but to use the Hadoop as a service. Yet, what firms like F2 can offer is their infrastructure for large scale projects where hardware and software has to be put together in fabric plants, mining and agriculture and where very diverse but extremely specific knowledge has to come together. Most importantly, the human resources can fly in from many different parts of the world to put together something new technologically and large in size.

One of the advantages of firms deploying Hadoop as a service is that firms do not need to worry about documentation. Documentation plays an important role in the advancement of the Apache Hadoop as discussed in the open-source literature. Furthermore, vendors do a great job maintaining the documentation of the Apache Hadoop up-to-date. The subject is covered under subjection 2.3.1. Free and Open-Source Software.

F13 is an organisation that combines big data (Apache Hadoop), artificial intelligence and machine learning to better understand human behaviour. It is a start-up and consists of professionals of very diverse but complimentary backgrounds – neuroscientists, anthropologists, sociologists, data scientists and software developers. F13's chief data scientist was keen to explain how important comprehensive documentation is for the organisation. The participant has pointed out that before deciding on distribution the organisation went through an extended evaluation process. One of the most important aspects of the decision to operate with the enterprise distribution was the quality and how vendors organise their documentation. As the chief data scientist says:

> *"The fact of having good documentation in both versions, and having a company to provide support and training in your country are very important things. The documentation has to be comprehensive and detailed. Most importantly, it should involve configuration and application. The documentation is essential for internal training."*

6.5.3. Conclusion

Multinationals expand its service offering through acquisitions as well as internal development. Whereas SLA is widely seen in outbound OI processes of vendors, multinationals can evade SLA through outbound the Apache Hadoop 'as a service'. It is since multinationals do not have the same level of source code governance as vendors. Start-ups are very likely to use freely available resources (community distribution and open data) to solve real-world problems.

Consultancies are keen to offer solutions to clients in advance for free before contracts are officially signed. Multinationals have developed a so-called Hadoop Lab for a monthly fee. It is an additionalS way to out the technology and commercialises the Apace Hadoop. Firms in the public sector and manufacture are more use to cloud-computing than telecommunication, banking and health sectors. Multinationals are now focussing on manufacturing because most data problems in the telecommunication and finance sectors have already been solved. Additionally, multinationals are more adaptable for large scale projects, e.g., mining, agriculture and manufacturing.

Bureaucracy dominates the public sector in Brazil. It makes it hard to innovate. It is one possible reason why the public sector lies far behind the private sector. The combination of data from the public and private sector can be used to address public health or other more general issues. Transparency is essential for data-centric organisations. Clients want to how the result was achieved. The rules of calculation, the so-called algorithms, are as crucial as the Apache Hadoop and the results achieved. Because, in the big data world, everyone uses the same ecosystem, competitive advantage can be gained with the rules of calculations.

CHAPTER 7. ANALYSIS & DISCUSSION

7.1. INTRODUCTION

The study pursued two sets of research objectives. First, it aimed to identify the contexts under which Apache Hadoop creators and consumers generate innovation streams and synthesise the outbound OI processes in firms' current markets. The research questions associated with the objectives include:

How do Apache Hadoop creators and consumers generate innovation streams in their current markets?

Secondly, the study sought to unravel the underlying conditions under which the Apache Hadoop evolves from a community base to enterprise platforms, and synthesise the outbound OI processes of Apache Hadoop creators and consumers in their current markets. The research questions associated with this objective include:

How can innovation streams evolve from community base to enterprise platforms?

To answer the two research questions, the thesis investigated the Apache Hadoop in eight industries in four case studies and presented the findings in Chapter 6.

We answer the two research questions by organising the findings. The chapter continues with a discussion of the theoretical and practical contributions of the study.

7.2. DISCUSSING OF KEY FINDINGS

7.2.1. How do Apache Hadoop Vendors generate New Innovation Streams?

According to Smith and Tushman (2005, p. 523), "an innovation stream refers to the portfolio of products simultaneously managed by an organization or strategic business unit". Products in this portfolio are defined relative to the technology and the target markets of the firm's existing product (Abernathy and Clark, 1985). For Tushman et al., (2020), successful businesses expand via consistent product and service development and through experimenting with various sorts of innovation. In comparison to the existing product, the firm's innovation can be incremental (Christensen, 1997; Dosi 1982; Shi et al., 2020), architectural (Henderson and Clark, 1991, Park and Tangpong, 2019), or discontinuous (Gatignon et al., 2002, Picaud-Bello et al., 2019).

The thesis applies innovation streams to analyse how creators and consumers of a digital ecosystem externalise their knowledge and commercialise their technology from a technology change viewpoint. According to Anderson and Tushman (1990), discontinuity is linked to a distinct approach to product and service design, and discontinuity might change the underlying mechanism of the goods and services. Outbound OI is reliant on dynamic capabilities. Therefore, having technology capabilities to enable outbound OI is crucial (Appleyard and Chesbrough, 2017). As a result, dynamic capacities, outbound OI, and ambidexterity relationships are mutually beneficial (Chesbrough et al., 2018).

Businesses that are ambidextrous can better adapt to changing business models and technologies (Birkinshaw et al., 2016; Hill and Birkinshaw, 2014). Ambidexterity in organisations arises from the successful mobilisation and integration of acts or behaviours to develop innovative configurations (Jansen et al., 2009; O'Reilly and Tushman, 2008).

Ambidexterity is likely to be applied at several organisational levels (Kassotaki et al., 2019), leading academics to consider the importance of multi-level analysis (Jansen et al., 2012; Raisch et al., 2009), (in Tarba et al., 2020). Effective organisations in dynamic environments are ambidextrous, capable of addressing current demands while adapting to future changes (Derbyshire, 2014; Enkel et al., 2017; Tushman and O'Reilly, 1996). Although helpful, ambidextrous organisations often face internal conflicts between two strategies, such as efficiency and flexibility (Adler et al., 1999) or alignment and adaptability (Andriopoulos and Lewis, 2009; Gibson and Birkinshaw, 2004) (in Lin and Ho, 2021).

Firm dynamic capabilities are crucial in gaining and retaining competitive advantage and developing and implementing a plan (Grant, 1991). Unique strategic and organisational actions may be defined as dynamic capabilities (Eisenhardt and Martin, 2000). According to Guerra et al. (2016), being ambidextrous is a means for organisations to build dynamic capacity. According to O'Reilly and Tushman (2008), ambidexterity is a dynamic capacity that increases a company's performance in changing contexts (Popadiuk et al., 2018). In other words, when exploitation and exploration are seen as dynamic capacities, they become helpful indicators for better understanding digital innovation ecosystems and their development.

Dynamic capabilities (outbound OI and ambidexterity) are a criterion for determining how businesses establish and retain a competitive edge in an ever-changing market. The concept of dynamic capabilities (explore new knowledge and exploit existing knowledge) will be borrowed from the framework of innovation streams to extend current studies on outbound OI by describing and illustrating some conditions under which ambidexterity can help organisations externalise their knowledge and commercialise their technology. Innovation stream is not used from the traditional organisational design perspective. But instead, the concept was borrowed in the study to illustrate how technology can evolve. By concentrating on firms' dynamic capacity to utilise and explore their knowledge and technology, this research emphasises the importance of ambidexterity in outbound OI operations.

The two study topics investigate how outbound OI and ambidexterity (dynamic capabilities) may be used to understand better how Apache Hadoop developers and consumers externalise their knowledge and commercialise their ideas in today's marketplaces. Based on innovation streams (Tushman and Smith, 2002; Benner and Tushman, 2003; O'Reilly and Tushman, 2008; Tushman et al., 2010), this research looks at how Apache Hadoop creators and consumers use their outbound OI to exploit and explore their knowledge and technologies and how these dynamic capabilities affect the ecosystem's evolution over time. This thesis aims to show how the move from exploitation to exploration and vice versa (from discontinuity to a dominating design) is crucial for Apache Hadoop's evolutionary system, and innovation streams provide a theoretical prism through which this might be explained.

The analysis suggests that Apache Hadoop vendors are ambidextrous by design. Vendors offer the Apache Hadoop as community and enterprise distributions. The community distribution is free, often consists of newly released codes, and needs much editing. The enterprise distribution is made up of matured codes that have been edited over a long period. While Apache Hadoop vendors explore new horizons through community distribution, they exploit the Apache Hadoop through enterprise distribution.

7.2.1.1. Discontinuous Innovation

The study found that internal data problems occur when the process of technological change and innovation begins. Discontinuous innovation deals with radically new products and services (Meyers et al., 1989), and in this particular scenario, it refers to what Veryzer

(1998) specifies as ‘technological capabilities’. Additionally, firms must be adaptable and sensitive to discontinuous advances since changes are difficult to predict (McKelvey, 2004). Discontinuous innovations are new technologies, products, or business models that dramatically differ from a particular sector's current state of the art (Birkinshaw, 2007). Therefore, organisations must form and sustain inter-organisational relationships with various external stakeholders, including suppliers, in implementing discontinuous innovations (Picaud-Bello et al., 2019).

The integration of a complex platform such as the Apache Hadoop in firms’ IT infrastructure requires a lot of fine-tuning and testing. During organisational learning, firms experience many bugs and experience many code corrections. At the early stages of the implementation, the Apache Hadoop requires a lot of flexibility from the ecosystem at the software coding level. It is crucial to keep in mind the Apache Hadoop consists of thousands of free and open-source software components that have to be aligned together and work as one with the firms’ infrastructures to answer to specific business problems. Another critical factor is that firms usually acquire data from many different places and file formats.

In most cases, firms already have an infrastructure and workflows in place that have to be aligned with the new technology. The case studies indicate that it takes a lot of trial and error to achieve a satisfactory result and testing is critical for best performance. The study showed that most projects start with community distribution, and community distribution is often used to present results to clients or internally where no dedicated funds are available. This is often the case within consultancies, SMEs or internal projects, and multinational and financial institutions. The financial sector uses community distribution exclusively for ad-hoc projects or testing of new workflows.

Community distribution is the foundation of the Apache Hadoop. In addition to the firm level where the Apache Hadoop has to be integrated into firms' IT infrastructures, community distribution plays a critical role in the organic development of the Apache Hadoop. Committers release new components at a much faster rate at the community distribution level. This, in turn, allows the enterprise distribution to become mature over time. Thus, community distribution is always a few releases ahead of enterprise distribution and drives, in most cases, discontinuous innovation within the ecosystem and in firms. However, community distribution is very labour intensive.

Community distribution is more suitable for discontinuous innovation because it offers flexibility for firms to work around an implementation problem. Another exciting feature of community distribution is that it enables firms to try out new workflows very cost-efficiently. The study has also found that SMEs and start-ups regularly take advantage of community distribution to develop technical competencies. These types of firms develop particular skills and enter the market with practical business solutions. The chapter has a dedicated section (7.3. Apache Hadoop Vendors Outbound OI Processes) covering the issue. There is evidence showing that organisation size and the flexibility of community distribution enables firms to explore the Apache Hadoop very competitively. However, this depends heavily on the skills of the individuals in the organisations. It is crucial to keep in mind that many SMEs and start-ups had their start at a multinational or at an Apache Hadoop vendor. The mobility of human resources has not only to do with the entrepreneurialism but also with the scarcity of professionals.

7.2.1.2. Architectural Innovation

Competitor or supplier-driven innovation may threaten a firm's market position (Christensen, 1997; Wagner and Bode, 2014). Furthermore, external innovation may cause companies to reassess and eventually modify their make-buy sourcing strategy (Argyres et al., 2019; Bigelow, 2019). As the product's component links become more modular and standardised (Baldwin and Clark, 2000), firms no longer require custom-designed components, reducing the need for transaction cost management (Christensen et al., 2002), and the overall cost reduction dominates the competitive environment (Park and Tangpong, 2019). Henderson and Clark (1990, p. 10) characterise architectural innovation as “innovation that changes how the components of a product are linked together while leaving the core design concepts (and thus the basic knowledge of underlying the components) untouched”. To understand the critical role played by architectural innovation in the contexts under which Apache Hadoop vendors generate new innovations streams, it is essential to acknowledge the evolutionary dynamics of innovation streams. According to Tushman and O’Reilly (1997) and Tushman and Smith (2002), innovation streams can be used as a tool to illustrate how technology may evolve. Thus, architectural innovation plays a critical role in the trial-and-error processes within the Apache Hadoop until a dominant design materialises – the enterprise distribution.

The study has shown that there is a process of code migration from community to enterprise distribution. It is also understood that the Apache Hadoop and its components undergo an evolutionary process of maturation that is characterised by new components being released at a much faster pace at the community distribution. The process, in turn, results in a trade-off scenario where the Apache Hadoop offers certain flexibility because the code base is still not fully matured. Eventually, however, the Apache Hadoop also displays spans of variations (eras of ferment) that are characterised by a certain level of operability. Depending on how disruptive ideas develop compared to prior breakthroughs, they have a varying influence on markets. Ideas based on competitive substitution of new technology for old technology may explain technological progress in this setting (Fisher and Pry, 1971; Sahal, 1981; Utterback, Pistorius, and Yilmaz, 2020). According to Berg, Wustmans, and Bröring (2019), the dynamic competition between technologies may result in the dominance of new technology over older technologies (Lin and Ho, 2021). Because the code base of the Apache Hadoop develops into a much more stable architecture, the process of trial and error will endure within the community distribution until an enterprise distribution materialises as a fusion of several proven solutions (Utterback and Abernathy, 1975; Abernathy, 1978).

While the process can be explained from a rather abstract angle, in practice it means that organisations have the flexibility to combine different pieces of the Apache Hadoop ecosystem with a certain level of stability. This is due to the increased operability of an architecture of the Apache Hadoop that has matured over time in its path to community distribution – the dominant design.

7.2.1.3. Incremental Innovation

Incremental innovation is usually an addition to the current products/service portfolios of firms and follows a logical technical progression by extending processes and/or designs of products and/or services. Incremental innovation suggests small changes, exploits already confirmed designs, and usually strengthens the authority of organisations (Nelson and Winter, 1982; Ettlie et al., 1984; Dewar and Dutton, 1986; Tushman and Anderson, 1986). The success of organisational innovation is also built on incremental innovation (Oerlemans et al., 2013). Incremental innovation involves improving existing technology in form, function, and features to meet current customer desires (Margaret and Nathaniel, 2019). Incremental innovations may help organisations gain a competitive advantage by assisting firms to adapt to changing conditions (Un, 2010). Unfortunately, firms that undervalue incremental gains may experience deadly or near-fatal setbacks (Shi et al., 2020).

While community distribution is responsible for the organic evolution of the Apache Hadoop ecosystems, enterprise distribution contributes to the Apache Hadoop in terms of software codes and documentation. Although the study was not able to identify the exact criteria that are taken into consideration for the migration of codes from one distribution to the other, documentation seems to take centre stage. It is necessary to note that documentation is one of the most important tools in new implementation or fine-tuning. Documentation is where architecture and software engineers can look for possible solutions for problems.

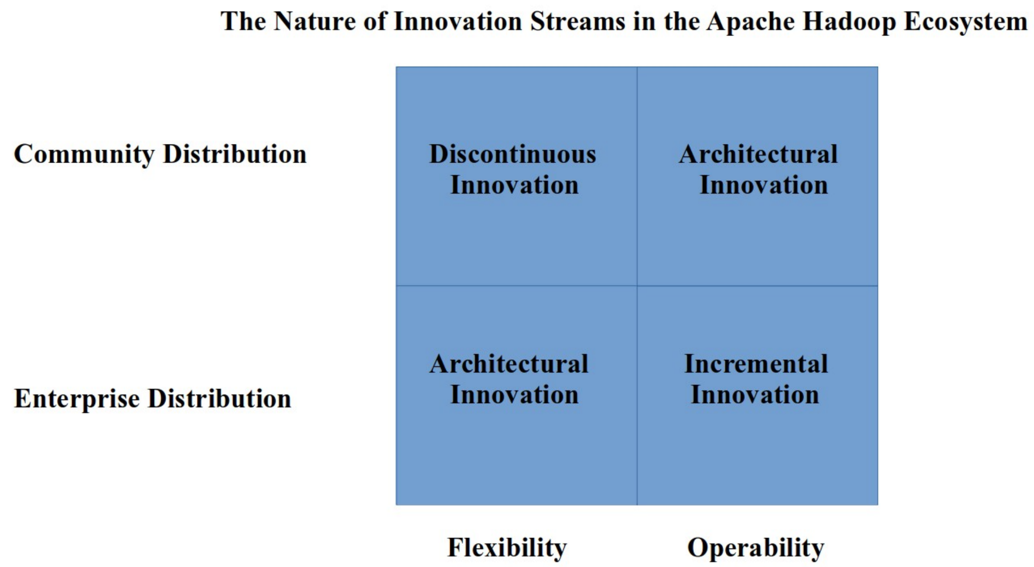
The study also suggests that while community distribution is often developed by SMEs and start-ups to solve very specific business problems, enterprise distribution covers a wide range of solutions, which sometimes makes it unaffordable for certain firms. Financial institutions that operate under strict compliance only run enterprise distribution due to the high fines associated with technical failures. The study points out that enterprise distribution offers more security in technical and juridical terms and it is extremely important to prevent issues with authorities and keep the image of the entity intact.

7.2.1.4. Ambidexterity in the Apache Hadoop Ecosystem

After identifying the contexts under which Apache Hadoop vendors generate new innovation streams, this section aims to illustrate them in a two-by-two matrix. Organisational ambidexterity is defined as the firm’s ability to explore new ideas and exploit current knowledge simultaneously (Andriopoulos and Lewis, 2009). One way to organise ‘exploration and exploitation’ in products and services development is the concept of innovation streams (Tushman and Smith, 2002; Benner and Tushman, 2003; O’Reilly and Tushman, 2008; Tushman et al., 2010). Innovation streams describe the evolving patterns of innovation (Tushman et al., 2020) and consist of continuous incremental innovation in existing products and services in conjunction with at least one non-incremental innovation (Smith and Tushman, 2005).

The study investigated the ways firms manage their innovation streams. The firms organised one or more type of innovation as they exploited and explored the Apache Hadoop. The innovations were distributed throughout different sectors and presented in four case studies. The study investigated the ability of firms to use the Apache Hadoop to innovate discontinuously, architecturally, and incrementally. The data suggest that community distribution was the onset of discontinuous innovation. For example, the Apache Hadoop ecosystem was the game-changer in terms of data management and analysis in firms across case studies and sectors. It has radically changed organisational thinking, and the study was able to replicate the discontinuity across the sectors and case studies. It appears that community distribution was critical for discontinuous innovation. The study suggests that organisations went through various test phases so that Apache Hadoop specialists could edit codes to achieve the best performance and align the Apache Hadoop with firms' information and technologies (IT) settings. Figure 8 illustrates the nature of innovation streams in the Apace Hadoop.

Figure 09. *The Nature of Innovation Streams in the Apache Hadoop Ecosystem*



The study found that some firms had to invest in new hardware and specialised human resources to be able to adapt to the latest technical requirements and to go through the learning curve and organisational change. What the study suggests is that in some cases, the discontinuity was so profound that although the organisation had the human resources and the know-how, they had to pay for a licence to be able to access the code. The community distribution creates a unique opportunity for organisational learning, technological change and innovation.

The study has also revealed that the flexibility of community distribution can help firms to solve problems with a technology that was developed to address a different issue. Firms that favour community distribution pursue flexibility instead of operability.

The final analysis suggests that firms that need to innovate discontinuously need the unparalleled flexibility of community distribution. Community distribution offers the resilience required by firms in the process of alignment of the firms' IT infrastructure with discontinuous technology (Apache Hadoop). In addition to the new technological asset, firms need to change workflow and organisational thinking to innovate discontinuously and to make data-driven decisions. The top left corner of the two-by-two matrix illustrates the flexibility of community distribution as it offers a more favourable environment for discontinuous innovation. This is synthesised in detail in section 7.2.1.1.

The top right corner of the two-by-two matrix depicts a platform that has become increasingly operational over the year via constant improvement, variation and selection. It is in line with studies on technological cycles by Tushman and Murmann (1998). According to Campbell (1969), for variation and selection to be consolidated in an evolutionary process, a retention mechanism is needed. The author believes that a successful variation needs to be perpetuated and reproduced.

The study suggests that in the evolving ecosystem of the Apache Hadoop, a selection period eventually starts once the code of the infrastructure demonstrates initial signs of maturation. The study found that there is no significant difference between community and enterprise distribution. It is understood that enterprise distribution offers some additional security features, such as better user-interface and compliance. When it comes to the core of the Apache Hadoop, however, both distributions are identical. The study also revealed that the release used in enterprise distribution is precisely the same one as the recommended community distribution software release. Vendors put a lot of effort into fine-tuning their enterprise distribution to perform meticulously. Although the recommended community distribution cannot confirm the effort, the issue will be clarified when the work mobilises Apache Hadoop vendors' outbound OI processes in section 7.3 to fill, at least partially, the research gaps presented in section 2.4.7. Research Gaps in Outbound OI.

Once the initial maturation phase crystallises, the study suggests four major implications: (1) the environment offers firms a great deal of integration with other Apache Hadoop components; (2) The study has also revealed that ad-hoc projects and testing often occurs with community distribution; (3) The analysis also suggests that there is a code migration from community to enterprise distribution, and the final analysis suggests that firms can migrate the codes tested by community to enterprise distribution almost effortlessly; (4) The final implication is that the bottom left and top right distributions are very much the same and integrate very comfortably with other components of the Apache Hadoop ecosystems, which is in line with the definition of Henderson and Clark. Architectural innovation is "innovation that changes how the components of a product are linked together while leaving the core design concepts (and thus the basic knowledge underlying the components) untouched" (Henderson and Clark, 1990, p. 10). The dominant design dictates and drives technical progress. Following the introduction of a dominant design, the technological trajectory may alter dramatically. Dominant design is characterised as the tipping point for opening industrial innovations via standardisation, allowing manufacturers to establish mass-manufacturing systems (Abernathy and Utterback, 1978; Murmann and Frenken, 2006). The dominant design in product development is validated by comparing technological trajectories before and after its introduction (Kim et al., 2021). At the stage of the technological cycle, the Apache Hadoop offers firms a trade-off. While the community distribution at the top right corner offer firms some flexibility and some operability, the enterprise distribution at the bottom left offers some flexibility, some operability, and compliance.

Finally, the bottom right illustrates the operational, yet slow and incrementally evolving enterprise distribution. Dominant designs and ensuing technological discontinuities characterise technology cycles (Anderson and Tushman, 1990; Van de Ven and Garud, 1993).

The success of organisational innovation is also built on incremental innovation (Oerlemans et al., 2013). Incremental innovation involves improving existing technology in form, function, and features to meet current customer desires (Margaret and Nathaniel, 2019). Incremental innovations may help organisations gain a competitive advantage by assisting firms to adapt to changing conditions (Un, 2010). Unfortunately, firms that undervalue incremental gains may experience deadly or near-fatal setbacks (Shi et al., 2020). Dominant designs meet the needs of wide-ranging users. Most importantly, it is not a radical innovation but rather a creative combination of prior innovations. Once a dominant design crystallises, further technological advancements constitute incremental improvements. The enhancements expand the standard and the technical system; the system grows into something more stable as one design expresses its dominance. The concept of dominant design has been used by several scholars to explain technological evolution (Anderson and Tushman, 1990).

7.2.2 How can Innovation Streams evolve from a Community Base to Enterprise Platforms?

After referring to the empirical data in the literature review and using the framework of innovation streams as a theoretical lens, the last section has met the first set of objectives by presenting the answers to the first research question and illustrating them in a two-by-two graph. The next section aims at completing the second set of research objectives by answering the research question and illustrating the technical and business cycle.

Similarly, to Section 7.2.1., Section 7.2.2. assumes that innovation consists of an evolutionary system characterised by a dominant design, issuing technological discontinuities (Anderson and Tushman, 1990; Van de Ven and Garud, 1993) and complex interactions between the different aspects of the innovation process (Edquist and Hommen, 1999).

7.2.2.1 Discontinuous Innovation

Internal Data Problems: As covered earlier in the discussion on the first set of objectives, the study found that technological change and innovation in firms deploying the Apache Hadoop starts with internal data problems in firms. It is with the community distribution that the firms' engineers edit codes and tune the Apache Hadoop for better performance - the beginning of the firms' journey through the deep blue sea of big data. The study has learned that the community distribution offers flexibility over operability. This condition provides an excellent environment for discontinuous innovation. Most importantly, the new technology has to be implemented in firms where technologies and workflows are already in place. Therefore, in the process of innovation and organisational learning, a comprehensive test phase is necessary. The study has learned how two firms have paid for a licence but prefer to deploy the community distribution due to its flexible qualities, even after the two firms went through learning curves. What makes the example so compelling for the analysis is the ability of Apache Hadoop vendors to be discontinuous even if they have to rewrite the whole code base of a module of the Apache Hadoop ecosystem from scratch. The study has also found that often consultancies, SMEs and starts-up are very likely to sell analytics services with prototypes developed in the community distribution. The advantage is that they can sell the very Apache Hadoop they have developed internally with little to no commitment by clients.

Development of Processes and Procedure: To innovate discontinuously, firms need to explore new ideas. In this respect, the case-studies indicate that firms without specialised human resources can innovate discontinuously by deploying cloud computing. This, in turn, results in major technological transformation within those firms. However, the development of processes and procedures goes far beyond the technical aspects. The study has also demonstrated that some firms had to rethink the way they calculate their investment return, which has a direct impact on the technical team running the Apache Hadoop within those firms. The size of the investment will dictate what hardware to buy and the human resources available to deploy and install the new technology. In terms of discontinuous innovation for start-ups with no technical capacities among their employees, cloud computing enables them to develop an algorithm and share the technological responsibility with a Hadoop-as-a-Service (HaaS) vendor. This is a case of firms having a competitive capacity with data but no Apache Hadoop skills.

7.2.2.2. Architectural Innovation

Portfolio of Solutions: According to Anderson and Tushman (1990), the dominant design develops out of an evolutionary process defined by variations and selection. The process eventually leads to a retention period. Firms' portfolios of solutions can be symbolised by the retention period. For instance, the study has shown that most features and modules can be found in both types of distribution (community and enterprise). The study also ascertained that during the process of implementation, firms go through a lot of testing to fine-tune and integrate the technology. Over time, the Apache becomes increasingly mature and can easily combine with other components of the Apache Hadoop that have gone through the same process of variation, selection and retention. The study shows that such firms are very likely to deploy Hadoop-as-a-Service. While these firms will be able to innovate discontinuously in mathematical terms in the form of innovative algorithms, they will be constrained by the Apache Hadoop in terms of discontinuity.

Another example of the process of variation, selection and retention was provided by firms that have gone through the learning curve and have managed to implement a running

Apache Hadoop. The study has shown that it is difficult to add a new component in a running Apache Hadoop framework, much like changing a flat tyre on a running car. Firms often integrate the new component on a community distribution, and when they achieve a satisfactory result, they migrate the code to the running Apache Hadoop. The study has shown that interruption in service is more than just technical failures: it has a potentially significant impact on the financial performance of firms; it can go as far as legal consequences and can damage the reputation of firms.

Service Offering: Once a firm has developed a stable Apache Hadoop platform, they can start offering services to clients. The study indicates that firms often go from internal problems to developing skills that are in high demand in the big data sector. The service offering is linked with the outbound capacities of firms and will be discussed in more details in Section 7.3. Apache Hadoop Vendors Outbound OI Processes.

Own Platform: Although the chronological position of service offering and own platform are interchangeable, they are both critical parts of the evolutionary system of the Apache Hadoop ecosystem. The study has shown that some firms have first developed their platform before starting to deliver service to other firms, as was the case with some start-ups and SMEs. At the same time, more prominent firms with a portfolio of clients would offer service to their clients. The study has shown that in this regard, there is no right or wrong approach. It is crucial to keep in mind that ecosystems such as the Apache Hadoop undergo constant changes as interactions never rest (Kallinikos et al., 2013). It is also important to distinguish at this point that offering services with a running Apache Hadoop is very different from developing an Apache Hadoop that answers to very specific business problems.

7.2.2.3. Incremental Innovation

Enterprise Distribution: The study indicates that enterprise distribution is associated with the same stables codes that have been organised by the platform providers over many years. Those firms have gone through discontinuous and architectural streams of innovation and compliance processes in one of the significant globally recognised technology auditing firms. Enterprise distribution owners usually have credentials such as ISO. Most importantly, they fully understand the code of the ecosystem. The study has found that the level of governance of the Apache Hadoop enterprise distribution owners can differ. Therefore, multinationals such as IBM sell the enterprise distribution but only as Hadoop-as-a-Service (HaaS). Because they do not have 100% control of the codes, they can overcome compliance by offering the Apache Hadoop as a service.

Service-Level Agreement: Enterprise distribution with a service-level agreement offers all the features of an enterprise distribution. The only difference is that the level of service is provided by firms that govern the codes fully. The significant difference between the type of vendors and multinationals like IBM is that the source code of the Apache Hadoop is open and the Apache Hadoop is not offered as cloud computing (HaaS). These are the firms that, in most cases, helped the multinationals to develop their platforms in the first place. Usually, these firms are the main contributors to the Apache Hadoop ecosystems, e.g. Cloudera and DataStax.

7.3. Apache Hadoop Vendors' outbound OI Processes

After answering both research questions and illustrating how Apache Hadoop vendors generate innovation streams in a two-by-two matrix and how the Apache evolves from a community base to enterprise platforms in a technological and business graph, this section aims at completing the two sets of research objectives by synthesising the outbound OI processes of Apache Hadoop creators and consumers in their current markets. Additionally, this section also aims at filling the suggested literature gap, at least partially. First, this section revisits the three research questions.

Literature Gap 1:

Outbound OI is usually considered from the perspective of the technology creator, and mostly from an IP-licensing angle. The thesis contributes to the OI literature by adding empirical evidence of outbound OI processes of technology consumers and creators in their current markets.

Literature Gap 2:

There are very few empirical studies to date on outbound OI processes in free and open-source software service digital platforms ecosystems. The work aims at adding empirical evidence to studies on service digital platforms ecosystems.

Literature Gap 3:

Although the literature on ambidexterity is very extensive, there is a lack in studies on ambidexterity in outbound OI and scholars have called for additional investigation. Therefore, the study answers to the recent call for more research on ambidexterity in outbound OI processes in firms.

To understand how Apache Hadoop creators and consumers commercialise the Apache Hadoop and outbound the technology in their current markets, it was important to learn how these firms organise their innovation streams. It is essential to consider that the Apache Hadoop consists of two types of distribution, and the platform undergoes a process of code maturation (from community to enterprise and from flexibility to efficiency).

The Apache Hadoop is an evolutionary system similar to the theories on technological cycles that are characterised by discontinuities that finish with the draft of a dominant design (Anderson and Tushman, 1990). Investment in big data skills and resources has risen in recent years (Gandomi and Haider, 2015; Goes, 2014; Gupta et al., 2018; Tambe, 2014; Yaqoob et al., 2016). Big data is especially intriguing as a digital innovation subject because it is disruptive and transformational throughout the information value chain (Abbasi et al., 2016; de Camargo Fiorini et al., 2018). Innovation in big data (like other innovations) has progressed (from business intelligence and data mining to data warehousing) by incorporating new resources and dynamic capabilities (Chae, 2019). Hence, discontinuity by exploring new ideas with the community distribution and the enterprise distribution is exploited incrementally. It is also in line with theories on digital ecosystems that consider digital platforms as always going through updates and changes. Additionally, to capitalise on the information technology (IT) infrastructure, organisations need to “explore and exploit” their information technology (IT) capabilities.

The processes of “exploration and exploitation” are in line with Levinthal and March’s (1993) ideas of how firms seek either exploitation for efficiency or exploration for innovation, or both at the same time. Finally, Apache Hadoop creators and consumers innovate discontinuously, architecturally and incrementally. Thus, it is of paramount importance to organise the innovation streams of these firms before the work could shed light on how technology creators and consumers commercialise the Apache Hadoop and outbound the technology in their current markets.

The study has found that community distribution is a critical component of external technology commercialisation. Furthermore, the study has demonstrated that firms across the four case studies develop solutions for clients and externalise their knowledge through community and enterprise distributions by selling consultancy, implementation, licensing and maintenance. This is consistent with the literature on outbound OI and free and open-source software. On the one hand, the literature of outbound OI has long been associated with IP-licensing. Other aspects of the commercialisation of technology assets were somewhat neglected or not well discussed. On the other hand, the traditional open-source software literature of business models is based on consulting, implementation and maintenance of the technologies. This section aims at adding other process models of technology commercialisation and external knowledge exploitation.

A result not consistent with the literature on outbound OI is the commercialisation of technology assets by technology consumers. Technology consumers are firms that do not directly develop the Apache Hadoop for external technology exploitation or even for internal usage. In other words, the Apache Hadoop does not belong to the primary business models of technology consumers, which makes the outbound OI process even more compelling for the study.

The study has found that a big data hardware seller has developed knowledge on the Apache Hadoop through selling their equipment over the years. It is important to note that hardware is a critical part of any IT infrastructure, including the Apache Hadoop. Through interactions with clients due to their data problems or hardware upgrades, the hardware seller was able to develop a portfolio of solutions. This is in line with firms that want to desorb knowledge to their partners (suppliers, customers and competitors). These firms need to make sure that the dialogue is harmonious with their respective partners' absorptive capacities, and even provide support with the task if necessary (Pagellet al., 2010; Lee et al., 2014). Yet, to some extent, the result is not consistent with the current literature on outbound OI. It is understood in the outbound OI literature that external technology exploitation is less practised because most firms are users instead of creators of technologies. Additionally, outbound OI is not necessarily associated with the strong appropriability of technologies (Christensen, 2006; West, 2006; West et al., 2006; Dahlander and Gann, 2010). The study has shown how an organisation that does not develop the Apache Hadoop internally or even work with it, has developed outbound practices to externalise its knowledge and commercialise its main product, which is hardware.

So far, much of the research on outbound OI has ignored the interdependencies of the activity of other firms in the ecosystems (Vanhaverbeke et al., 2014) and little progress has been made with analyses of outbound OI at the firm level (West et al., 2006; Chesbrough and Bogers, 2014). On the other hand, Masucci et al. (2020) have explored outbound OI strategies in ecosystems, bringing to light mechanisms used by firms to accelerate technological progress in their complementary activities (Masucci et al., 2020). The study conducted by Masucci et al. was about the oil and gas industry only, and there was no direct link to free and open-source software. In this regard, this study can add to the literature on outbound OI by bringing some empirical evidence to light of mechanisms of external technology commercialisation in a free and open-source software ecosystem.

It is understood in the OI literature that any inbound flow for an organisation should generate an outbound flow for another. Although it seems to be logical, firms are more likely to practise inbound over outbound OI. However, an example of inter-organisational collaboration in terms of inbound and outbound processes simultaneously is Apache Cassandra and Microsoft Azure. It is critical to mention that although neither of the two firms belongs to my data directly, an organisation that uses the Apache Cassandra through the Microsoft Azure cloud service does. The collaborative alliance provides the study with two main lessons: (1) the DataStax Enterprise (DSE) does not support cloud services, and

Microsoft Azure does. Therefore, DataStax can now outflow their existing knowledge and technology via Microsoft Azure cloud infrastructures. However, the DataStax alliance represents much more than outbound OI practices within its Apache Hadoop ecosystem. It suggests access to potentially all clients of Microsft Azure worldwide. The alliance provides DataStax with the possibility of externalising its technology assets without the need to go through compliance for cloud service. The alliance is a different process model for external knowledge exploitation because it does not only represent the commercialisation of technology assets on its own, it symbolises a firm overcoming compliance barriers through an alliance in its ecosystem. (2) At the other end of the spectrum, and consistent with the results of the study, Microsoft Azure can now go through the technological and business cycles and generate innovation streams to develop its own Apache Cassandra service by aligning it with the HaaS they already commercialise. In terms of cloud computing (HaaS), Microsoft Azure could profit from little resistance and immense trust from firms in the technology. The study has shown that cloud computing is widely accepted by participants.

A very different process model for external knowledge exploitation and technology commercialisation was another case of DataStax. DataStax decided to rewrite the Apache Cassandra 2.0. The implications of the simple of act of rewriting an entire module are huge. For instance, since DataStax governs the code base of Apache Cassandra, they can now dictate the direction of its development. In this regard, DataStax excelled in making the Apache Cassandra feel and look like a traditional SQL database. SQL is widely used by many firms across industries for data management. Without going into too many technicalities, structured data can be understood as the traditional data that consists of numbers and letters seen in most of the datasheets of organisations across the board. By making the Apache Cassandra operate and look like an SQL database, they made the transition of professionals much smoother. Consequently, they set the ground for the external commercialisation of its technology assets – the Apache Cassandra. The second implication is that the DataStax development team can now go through the technological and business cycles and develop the Apache Cassandra from community distribution to the enterprise platform with a service-level agreement.

A result consistent with the literature on open-source software is that the community distribution is always some releases ahead of the enterprise distribution. The study has shown that the progress of the ecosystems is dictated by community distribution. Yet, financial institutions operate under strong regulation by domestic and international bodies. The study has demonstrated that on different occasions, a component of the community distribution could be beneficial for the institution. Still, because the component is not at the enterprise distribution, the bank cannot use it. The type of outbound practice was found to be very difficult because it takes a lot of communication and the alignment of both firms. Most importantly, the integration has to take place with minimal service disruption. This type of result is not necessarily consistent with the literature on outbound OI within an ecosystem. There is little literature on outbound OI, demonstrating the challenges of external technology exploitation when compliance plays a critical role in the outflow of knowledge within the ecosystem such as the Apache Hadoop. The study has demonstrated that the type of external commercialisation of the existing knowledge and technology is very demanding and requires a lot of experience of both firms. The inbound flow of an organisation generates an outbound flow for another one. The study demonstrated that this innovation process requires a different level of trust between the two organisations. Although it seems to be a trivial process, it is a process that can be very risky due to platform failure or disruption in the service, which can cost millions for organisations as fines associated with technical failures are incredibly high.

The study has established that there are different levels of governance of the codes of the Apache Hadoop. The different levels of governance were illustrated in the technological and business cycle graph after answering the second research question. The study has demonstrated that the external commercialisation of knowledge and technology differs substantially depending on the level of governance of the Apache Hadoop source code. While vendors such as Cloudera and DataStax can offer service-level agreements, multinationals are more likely to offer HaaS. Although not consistent and little discussed in the literature of OI, the outbound practices of the two types of firms were found to be very different. While Cloudera and DataStax put in a lot of effort to govern the organic evolution from the community to the enterprise distribution within the Apache Hadoop ecosystem, multinationals are more likely to offer a cloud computing solution for clients. The size and the types of projects run by firms are little discussed and inconsistent with the literature on outbound practices. While Cloudera and DataStax exercise an absolute dominance in industries such as telecommunication, finance and marketing multinationals are gaining a lot of ground when it comes to agriculture and large-scale industrial production. The external technology as such requires very different types of logistics and the embedment of various other kinds of technologies simultaneously.

Additionally, multinationals are very used to handling such big projects. The study has also shown that most of the innovation within the telecommunication, financial and marketing sectors has already taken place when it comes to the Apache Hadoop but the Apache Hadoop is still in its infancy in high-scale industrial production and agriculture. Big data and machine learning has gained substantial importance in those industries over the last couple of years and it is expected to penetrate the sector even more.

A different process model for external knowledge exploitation is closely linked to firms in large-scale industrial production and agriculture. Consistent with the literature on outbound OI and this study is that firms use community distribution to show initial results with little financial commitment from clients. However, the study found that one of the significant HaaS providers in Brazil has developed a so-called ‘Hadoop Lab’. The Hadoop Lab is a way in which a multinational can externally exploit its technology and commercialise its existing knowledge. The study has learned that some projects can have a colossal magnitude, and the testing phase can take months before the clients have satisfactory results. The study even discussed an example in the iron industry in which the multinational and its clients had to test for months before the Apache Hadoop was aligned with the client’s set-up and some useful analytics were possible.

Consistent with the study is that firms usually fine-tune the Apache Hadoop and eventually present the results with little commitment to clients. This, in turn, enables firms to overcome their initial scepticism of clients and outbound the Apache Hadoop. The study found that the concept was often used by start-ups and SMEs very successfully. Some participants have even claimed that it has helped them to close many deals, and the model has had an enormous positive impact on the firms’ external commercialisation of existing knowledge and technology. The example of the multinational indicates, however, that the organisation has developed a “lucrative” way to commercialise the test phase and externalise the Apache Hadoop. With the Hadoop Lab, the multinational was not only able to pay for expenses during the long periods of testing and fine-tuning, it was also able to gain a reputation for cutting-edge technology in the Brazilian scenario. Most importantly, the multinational could gain valuable experience and knowledge in the process.

External technology knowledge exploitation as such was also observed in firms that depend on the Apache Hadoop for daily business but that are not very conversant with the technology in its full capacity, as was seen in many firms that specialise in data. The study has shown that firms deploying the Apache Hadoop need two types of professionals: data scientists and software engineers. Although some firms consist only of data scientists and some of the data scientists may have some knowledge of the Apache Hadoop and may even be able to install the ecosystem, when it comes to coding, they have their limitations. Therefore, these types of firms are very likely to use HaaS or enterprise distribution. As previously mentioned, their main business model comprises data and algorithms. Since algorithms are firms' primary revenue stream, they have been improved over many years. The Apache Hadoop is seen as a tool to manage data, calculate results, externalise knowledge and outbound firms' technology (algorithms). Most importantly, in most cases, certain types of firms would be unable to operate lucratively or even cost-effectively without the Apache Hadoop ecosystem.

CHAPTER 8. CONCLUSION AND IMPLICATIONS

8.1. SUMMARY OF MAIN FINDINGS

8.1.1. How do Apache Hadoop Vendors generate Innovation Streams?

Consistent with the innovation streams literature, the analysis suggests that the generation of innovation streams is characterised by evolutionary systems and consists of incremental, and at least one discontinuous, innovation. Yet, this study looks to confirm the assumption that Apache Hadoop vendors are ambidextrous by design. While they explore new horizons through community distribution, they exploit the Apache Hadoop through the servitisation of the community distribution, e.g., enterprise distribution and HaaS.

The study shows that discontinuity, technological change and innovation start with internal data problems. It is during a period of organisational learning that firms experience substantial code correction and bug fixing. While firms correct codes and debug errors, the Apache Hadoop is aligned with other Apache Hadoop components and the firms' IT assets. In this context, firms need to develop new work-flows and processes.

The study suggests that community distribution is the foundation of the Apache Hadoop at all levels. It plays a critical role in the organic development of the ecosystem as a whole. The study has shown that the Apache Hadoop and its components undergo an evolutionary process of code maturation that is characterised by new components being released at a much faster pace in community distribution. It takes a lot of trial and error to bring about satisfying outcomes with the Apache Hadoop. At an early stage, it calls for a great deal of flexibility from the ecosystem at the software coding level. Community distribution seems to be more suitable for discontinuity due to its flexible qualities which, in turn, enables firms to work around an implementation problem. The process of trial and error will endure with community distribution until an enterprise distribution materialises as a fusion of several proven solutions. The process is explained in abstract terms as spans of variation and retention mechanisms. Eventually, the process results in a trade-off scenario where the Apache Hadoop offers certain flexibility and a certain level of operability.

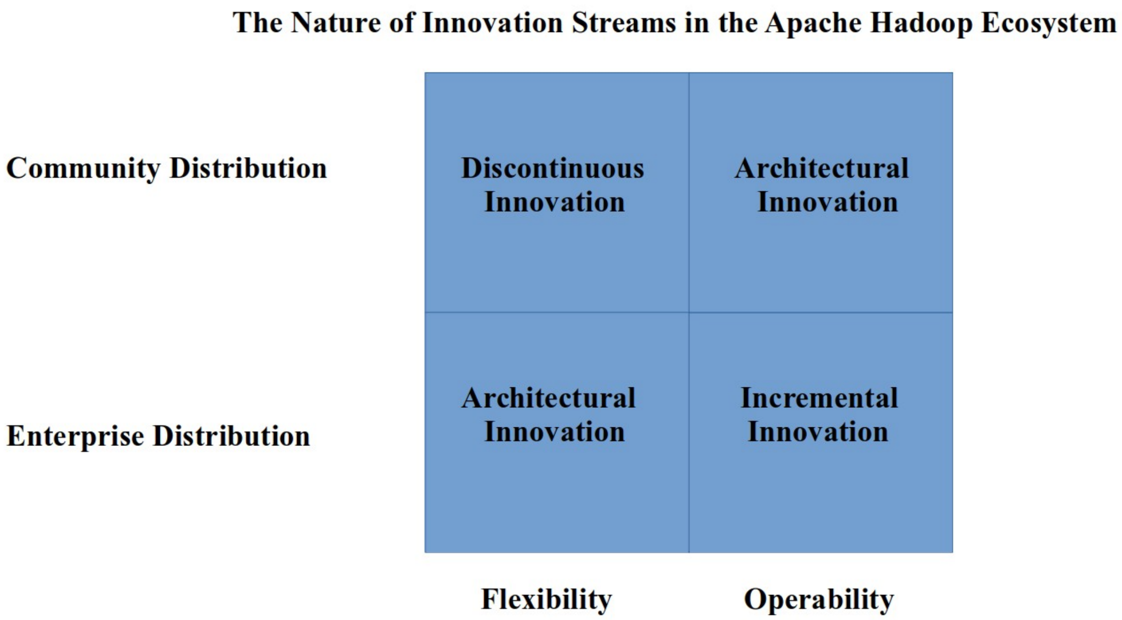
While community distribution is responsible for the organic evolution of the Apache Hadoop ecosystems, enterprise distribution contributes to this in terms of software codes, code syntax and documentation. What the study has found is that there is a constant code migration from community to enterprise distribution. But the study was not able to give a clear overview of the criteria under which codes are migrated from one distribution system to the other. Having said this, documentation seems to be one of the main pre-requisites.

The second research goal, and in conjunction with the first research question, was intended to illustrate the process of generation of the innovation streams in a two-by-two matrix. This is presented and explained again below in Figure 10.

The main purpose of the two-by-two matrix was to illustrate the ability of Apache Hadoop vendors to innovate discontinuously, architecturally and incrementally depending on which distribution (community or enterprise) was being used. What the study has found out is that there is a relationship between the way firms may innovate and the flexibility and operability of the distribution (community and enterprise) firms deploy.

The top right corner illustrates the flexibility of community distribution as it offers a more favourable environment for discontinuity. The top right corner depicts a platform that has become increasingly operational over the years via constant improvement, variations and selection. The bottom left and top right are very much the same distributions and integrate very comfortably with other components of the Apache Hadoop ecosystem, fitting very well in the definition of architectural innovation. The bottom right distribution meets the needs of wide-ranging users. Once a dominant design crystalizes, further technological progress consists of incremental improvements.

Figure 10. *The Nature of Innovation Streams in the Apache Hadoop*



8.1.2. How can Innovation Streams evolve from a Community Base to an Enterprise Platform?

As with the first research question, this section assumes an evolutionary system characterised by a dominant design issuing technological discontinuity and complex interactions between the different aspects of the innovation process. The study has identified seven significant phases the Apache Hadoop undergoes when moving from community distribution to enterprise distribution with a service-level agreement. The study has used the theoretical lens of innovation streams to explain the technical and business cycles of the Apache Hadoop, as illustrated in Figure 9.

Phase one is represented in the graphic as an internal data problem. It is where the process of discontinuity starts. Apache Hadoop engineers use community distribution to edit codes and tune the platform for better performance. Phase one can be described as the beginning of the firms' journey through the deep blue sea of big data. The flexibility of community distribution provides an excellent environment for discontinuous innovation.

The development of processes and workflow is defined by the need of firms to innovate discontinuously by exploring new ideas. The phase goes far beyond the technical aspect related to the alignment of the Apache Hadoop within firms. The study has demonstrated that firms have to rethink the way they operate as a whole; in some cases, even the way they calculate returns of investments. Most importantly, firms have to start making data-driven decisions, which is very different from existing processes in firms. At the other end of the spectrum, big data service providers innovate discontinuously by developing solutions for real business problems. The type of solutions they suggest are designed internally or following interactions with customers. To do this, firms usually have to make a significant investment in technology and human resources.

While the first two phases are rather explorative and are associated with the exploration of ideas new to firms, the third and fourth phases of the technical and business cycle are linked with the specific retention of the variations explored in the first two phases. Once a firm has developed a stable Apache Hadoop, they can offer a variety of service to clients. Firms that were capable of developing their own platforms are very intimate with the Apache Hadoop. There have been able to go through significant organisational changes very successfully and have accumulated solutions that they have refined over the years. The code of the Apache Hadoop has become mature in the process and is now only incrementally improved.

Finally, the Apache Hadoop reaches a level where the governance of the code plays a significant role. While some vendors with 100% governance of this code can offer enterprise distribution with a service level agreement, other firms that are very intimate with the Apache Hadoop develop their HaaS platforms. The study has found that HaaS vendors can overcome compliance by denying access to the source code to customers by offering the Apache Hadoop as a cloud service.

8.2.3. Outbound OI Processes

To complete the three sets of objectives associated with the two research questions, this section synthesises the outbound OI processes used by Apache creators and consumers in their current markets. By doing so, this study attempts to fill the literature gap, at least partially. Before suggesting the contribution to the outbound OI literature, the literature gaps are revisited.

Literature Gap 1:

Outbound OI is usually considered from the technology creator perspective, mostly from the IP-licensing angle. The thesis contributes to the OI literature by adding empirical evidence of the outbound OI processes of technology consumers and creators in the firms' various current markets. Additionally, recommendations were made to identify and analyse additional case studies.

Contribution to Literature Gap 1:

The study has shown that through the sales of the Apache Hadoop server over the years, a big data hardware seller has accumulated valuable knowledge concerning the Apache Hadoop. Hardware is a critical part of any IT infrastructure, including the Apache Hadoop. The hardware seller was able to build a portfolio of solutions through interactions with customers based on their data issues or hardware updates. It is in line with the need for companies who wish to desorb information to their partners (suppliers, customers and competitors) to ensure that the conversation is compatible with the absorptive capabilities of their respective partners, and also provide help with the task, if necessary (Pagell et al., 2010; Lee et al., 2014). However, the outcome is not compatible with the current literature on outbound OI. It is well accepted in the literature on outbound OI that external technology exploitation is less practised because most businesses are consumers rather than technology developers. Furthermore, outbound OI is not inherently related to the strong appropriability of technologies (Cristensen, 2006; West, 2006; West et al., 2006; Dahlander and Gann, 2010).

The study was able to demonstrate with empirical evidence how a company that has not developed, or even operated internally, the Apache Hadoop has successfully generated outbound OI practices to externalise its expertise and market its primary hardware products.

The thesis has discussed outbound OI processes of Apache Hadoop creators and consumers in eight industries in firms' current markets. The case histories section offered examples of outbound OI processes in more traditional industries, e.g. telecommunication, finance, retail, and marketing. Additionally, the study has also answered the call by different authors (Smith and Akram, 2017; West and Bogers, 2017; Bogers et al., 2018; Zhen et al., 2018; Mohamad Hashim et al., 2020) to identify and analyse outbound OI processes in the public sector.

Literature Gap 2:

There is little evidence of outbound OI processes in free and open-source software service digital platform ecosystems (Vanhaverbeke et al., 2014), and little progress has been made with analysing outbound OI at the firm level (West et al., 2006; Chesbrough and Bogers, 2014).

Contribution to Literature Gap 2:

Outbound open innovation research has so far overlooked the interdependencies with the activities of firms in ecosystems (Vanhaverbeke et al., 2014). Thus, this thesis has also answered more recent calls (Verreynne et al., 2020; Cheah and Ho, 2021) for research in outbound OI in digital service platforms. Outbound OI processes in ecosystems have been recently discussed by Masucci et al. (2020), revealing methods used by businesses to drive technological change and innovation in their complementary activities. This study was mainly about the oil and gas industries, and there was no direct connection to digital ecosystems. In this respect, by bringing some empirical evidence of external technology commercialisation processes in free and open-source software to light, this study adds to the literature of outbound OI.

An example of inter-organisational collaboration in terms of outbound OI processes in free and open-source software ecosystems is the alliance of DataStax (Apache Cassandra) and Microsoft (Microsoft Azure). It is important to note that neither of the two companies belongs directly to my data directly. But an organisation that uses Apache Cassandra via the cloud services of Microsoft Azure does. The collaborative alliance offers two main empirical contributions to the study.

First, DataStax Enterprise (DSE) does not support compliance for a cloud server, as opposed to Microsoft Azure. DataStax can now externalise their knowledge and commercialise the Apache Cassandra via Microsoft Azure. Most importantly, DataStax can, through the alliance with Microsoft, overcome compliance. For DataStax, the partnership represents much more than the outflow of knowledge and technology. It implies access to all Microsoft Azure customers worldwide. It is a different process model for external knowledge exploitation because it not only represents the commercialisation of technology assets on its own, it symbolises a firm overcoming compliance barriers through an alliance in its ecosystem.

Second, this is consistent with the results of the study yet not necessarily with the literature of outbound OI. Microsoft Azure can go through the technical and business cycle to evolve its own Apache Cassandra solution by aligning it with the HaaS they are already selling.

Apache Cassandra embodies another very different process model of knowledge exploitation and technology commercialisation in a free and open-source software ecosystem. DataStax was tasked to rewrite the Apache Cassandra 2.0 from scratch. This alone is rather unique, and the effects of the simple act of rewriting an entire module were colossal for the ecosystem. Since DataStax governed the development of Apache Cassandra, they could dictate the direction of its development. In this respect, SQL is a commonly used database technology by many organisations across industries, and DataStax has done an outstanding job in making the Apache Cassandra look and feel like a typical SQL database. By doing so, DataStax made the transition of database professionals to the Apache Cassandra much smoother. Additionally, the generic SQL look and feel made the decision of firms to invest in the module straightforward because it offered a “similar” technology to the one they had been working with for a long time. The database is associated with less training for employees and with shorter organisational learning processes. Consequently, DataStax laid the foundation for external commercialisation of its technological asset. It can be said that Apache Cassandra is one of the most widespread database technologies available today.

The second implication in this regard is that firms that use the Apache Cassandra can employ SQL professionals. The study has learned that big data professionals are scarce worldwide. As the Apache Cassandra looks and feels like any other SQL database, firms using it can overcome the scarcity of qualified human resources by training their new employees in the Apache Cassandra’s workflows.

This study has found that community distribution dictates the direction of the development of enterprise distribution. Features that have found a business use in community distribution will eventually migrate to enterprise distribution. Although there are features that are exclusively developed for enterprise distribution (user-interface and security features), community distribution is always ahead of enterprise distribution in terms of new features. The configuration contributes to the organic development of the ecosystem and is consistent with the literature on free and open-source software. Also, well discussed in the outbound OI literature is that financial institutions work well under strong domestic and international regulation. On many occasions, a component of community distribution could be very beneficial for an organisation. Still, because the feature is not available to enterprise distribution, the organisation cannot use it.

This type of outbound activity was found to be very challenging since both organisations involved need a lot of experience, coordination and alignment. The integration of a new component must take place with minimum disruption to any services. There is little empirical evidence in the outbound OI literature covering the challenges of external technology exploitation when compliance plays a critical role in the outflow of knowledge within an ecosystem, such as the Apache Hadoop. What the study has revealed in this regard is that the two firms have to work very closely together to push the release of the feature. Both organisations must follow a tight schedule because any failure is very likely to result in a disruption to services.

There are unprecedented levels of code governance within the Apache Hadoop ecosystem. The external commercialisation of knowledge and technology contrasts significantly while depending on it. Whereas vendors (Cloudera and DataStax) offer service-level agreements, multinationals are more likely to offer HaaS. Although not consistent with and little discussed in the literature of outbound OI, it stands that the two sorts of firms were found to be exceptionally diverse in their outbound OI processes. Whereas Cloudera and DataStax focus on governing the organic evolution from community to enterprise distribution, multinationals are more likely to offer a cloud computing service to clients (Hewlett and Packard Enterprise and Microsoft Azure). Little examined and in conflict with the writing on outbound OI practices appears to be the size and types of the ventures run by one or the other. Cloudera and DataStax exercise an absolute dominance in telecommunications, finance and marketing; multinationals are gaining ground when it comes to agriculture and large-scale industrial production. External knowledge exploitation and technology commercialisation requires exceptionally diverse sorts of logistics and the embedding of various kinds of technologies simultaneously. With this in mind, multinationals are in better positions to take projects that have international dimensions.

Literature Gap 3:

Although the literature on ambidexterity is extensive, here the thesis answers to a recent call for more research on ambidexterity in outbound OI processes in firms. The essence and environments under which it may take place are still unclear (Drechsler and Natter, 2012) and have been widely neglected (Florén and Frishammar, 2012; Jia and Lamming, 2013; Scherrer-Rathje et al., 2014; Hu et al., 2015; Meinlschmidt et al., 2016). More recent studies (Dubrowska et al., 2019; Li et al., 2020) have also highlighted the lack of research in ambidexterity in outbound open innovation and the thesis answers to call for additional studies.

Contribution to Literature Gap 3:

The next contribution is directly linked to large-scale industrial production, agriculture and the way HaaS service providers explore and exploit the Apache Hadoop ecosystem. The study has shown that firms are very likely to use community distribution to display their initial outcomes, with little financial commitment from the clients. The possibility of delivering results to clients in advance had had a positive impact on many firms in the sample; it is a strategy used quite often by consultancies, start-ups and SMEs. The research, however, brought to light a so-called Hadoop Lab. The Hadoop Lab was developed and put into operation by one of the significant HaaS providers in Brazil. The Hadoop Lab is a very effective way for a global corporation to leverage (explore and exploit) its knowledge externally, and to commercialise the Apache Hadoop. The study found out that for projects of a specific scale, it can take months for the testing process to achieve an acceptable result for clients. The issue was discussed using an example in the iron industry where the testing period went on for over six months before any useful analytics were even possible. The example shows one possible way the HaaS service provider found to externalise knowledge and commercialise the Apache Hadoop in the testing period of a large-scale project. In addition to the lucrative exploitation and exploration of the Apache Hadoop in the form of a Hadoop Lab, the HaaS service provider was able to affirm its reputation for cutting-edge technology in the Brazilian scenario. Most importantly, during the process, the HaaS service provider obtained additional knowledge and skills.

8.2. THEORETICAL CONTRIBUTIONS AND PRACTICAL IMPLICATIONS

8.2.1. Theoretical Contribution

The previous section discussed the main findings of the study with relation to the research questions and objectives. In what follows, the sections unite the findings and compare them with the extant literature to discuss their significance. Table 61 demonstrates an overview of the theoretical implications.

This research focuses on a digital ecosystem, especially data processing frameworks (Apache Hadoop). In the context of open-source software development, several writers have highlighted dynamic capabilities in the OI paradigm. Outbound OI poses issues around business model selection and technical strategy and the evolution and management of such dynamic capabilities (Bogers et al., 2019). While using reasoning as a foundation, the unit of analysis considers the organisation's capacity to apply its outbound OI processes to drive technological progress and innovation. Outbound OI and ambidexterity are dynamic capabilities employed by Apache Hadoop creators and users to externalise their knowledge and commercialise their technology. Innovation streams and the innovation classes are not applied in organisation design but instead from a technology change perspective and are used as the guiding concept to understand how technology may evolve.

Table 60. An Overview of the Theoretical Implications of the Study

Contribution	
An enriched understanding of the outbound OI processes of technology consumers and creators in their various firms' current markets	- The mobilisation of the role played by the different classes of innovation in the evolutionary process of discontinuity (community distribution) to the development of the dominant design (enterprise distribution). - The organisation of processes of knowledge externalisation and technology commercialisation of Apache Hadoop creators and consumers in their various current markets. - The development of a two-by-two matrix illustrating how Apache Hadoop creators and consumers may generate innovation streams.
An enhanced understanding of outbound OI processes in free and open-source software service digital platform ecosystems in their various current markets	-The articulation and differentiation of the different phases of the evolutionary process (from community distribution to enterprise distribution with service-level agreement). - The organisation of processes of knowledge externalisation and technology commercialisation of Apache Hadoop creators and consumers in their various current markets. - The development of a technologies and business graphic depicting how the Apache Hadoop may evolve from community base to enterprise distribution.
An extended understanding on the ambidexterity of Apache Hadoop creators and consumers in their various current markets.	- The articulation of some of the essences and environments under which ambidexterity in outbound OI takes place for Apache Hadoop creators and consumers in their various current markets.

Businesses are moving their innovation focus away from internal research and development facilities toward facilities outside their firm's boundaries as technological progress and expert knowledge become more scattered. This work supports macro-level analysis for digital innovation, referring to a "digital innovation ecosystem" as a complex mix of technology, processes, ideas, commercial applications, enterprises, and institutional contexts. Big data is particularly fascinating as a topic for digital innovation because it is disruptive and transformative throughout the information value chain and comprises information, skills, ideas, organisations, and other social and institutional contexts.

The first contribution of this study relates to the diversity of outbound OI research. Previous studies on outbound OI have been mostly based on the technology creators' perspective or the IP-licensing angle in big organisations. Thus, existing research on outbound OI has largely ignored other types of such innovation processes. In addition, there are few examples of outbound OI practices by technology creators and consumers in their current markets.

This study fills this gap, at least partially, and the approach has been extended. This work has illustrated outbound OI processes in firms' various current markets. Such an approach helps to understand the underlying mechanisms of knowledge externalisation and technology commercialisation by technology creators and consumers in their various current markets. It views innovation streams from technological change perspective and understands outbound OI and ambidexterity as dynamic capabilities. Additionally, the study has exposed and synthesised the crucial roles played by the different innovation classes in the process of knowledge externalisation and technology commercialisation.

The second contribution of the study is that it provides an extended understanding of knowledge externalisation and technology commercialisation in digital ecosystems. The extant literature on outbound OI has neglected interdependencies within the activities of firms in ecosystems. Additionally, little progress has been made at the firm level. This thesis expands the outbound OI literature from the angle of outbound OI within a digital ecosystem based on free and open-source software. The findings improve the understanding of the interdependent relationship between alliances within firms in ecosystems and contribute to the knowledge domain of this area of research. There is limited prior research available on such collaboration from the perspective of knowledge externalisation and technology commercialisation. The present study suggests the existence of particular connections between external knowledge exploitation and technology commercialisation within firms in ecosystems, in particular, a direct joint effort to instigate knowledge exploitation and technology commercialisation through technology alliances within a digital ecosystem.

The third contribution also advances an enhanced understanding of outbound OI processes in service digital ecosystems. The thesis contributes to the literature by providing a more in-depth understanding of knowledge externalisation and technology commercialisation of firms. The findings show how firms promoted underutilised internal technology outside the organisation by combining old established technology with their new technology and business models. The internal technology commercialisation enabled the transition from old to new, making it more comfortable for professionals and firms in their current various markets.

Finally, the fourth contribution related to the literature of ambidexterity in outbound OI. Although the two frames of literature are said to be complementary, there is limited prior research available. This thesis offers a two-by-two matrix illustrating the relationship between the flexibility and operability of codes and the way firms may innovate.

Additionally, knowledge exploitation and technology commercialisation have been synthesised and presented according to their innovation classes.

8.2.2. Managerial Implications

The first implication for firms is that there are different ways of innovating with the Apache Hadoop system, and each distribution offers a more suitable environment for a specific type of innovation. On the one hand, managers that want to innovate discontinuously are much better off with the flexibility of community distribution. On the other hand, for managers that want to innovate incrementally, the operability of the enterprise distribution is more appropriate. Another practical implication is firms that go through the learning curve by applying their dynamic capabilities and deploying the community distribution are very likely to innovate discontinuously and architecturally in the future. In contrast, firms that deploy enterprise distribution from scratch will be minimal in terms of their use of discontinuity and architectural innovation. Firms have to evaluate if the Apache Hadoop is in the centre of their business model or if it is only a means to an end.

Firms that want to become more data-driven in their decision-making need to go far beyond the sole act of technological change. They need to change their organisational thinking completely. It does not make sense for a firm to invest millions in the Apache Hadoop if it is not considering the data results. Most importantly, firms need to be able to ask the right questions. Therefore, it is not only about increasing the technical capacity of firms by organising a number of Apache Hadoop specialists to work together. The Apache Hadoop specialist is only one piece of the puzzle. To make data-driven decisions, firms need data engineers to work very closely together with Apache Hadoop engineers. In other words, the Apache Hadoop is only as good as the firms' data and Apache Hadoop engineer teams.

The third implication is that once firms have gone through the learning curve, they need to find ways to help other firms in the current market to implement and use the Apache Hadoop. The study has shown how a hardware firm was able to externalise its knowledge and sell their hardware in a package. Another firm was able to outbound their knowledge via an Apache Hadoop course and, at the same time, overcome the lack of adequate professionals in the country by scouting for the best course graduates. By externalising its knowledge, the firm acquired substantial financial and non-financial gains. Today, the firm is one of the most prominent and most profitable Apache Hadoop vendors in Brazil.

The study also learned that there was no university degree for data engineering in Brazil at the time of the fieldwork. Therefore, firms need to develop internal ways to train talented candidates from other industries, such as physics and engineering, in the world of the Apache Hadoop. It is a way to overcome the scarcity of labour in the country until the academic institutions fill the gap. Therefore, the use of the right coding language is critical. While there are many talented programmers, they are just not aware that their skills can be used in the data sector as well. Therefore, the right computing language is not only critical to the development of the right workflow but also for access to the right human resources.

This thesis offers managers insights into outbound OI practices as dynamic capability. It recommends that managers reassess the role of innovation within and beyond organisational boundaries: most importantly, the role of innovation in peer-to-peer collaboration in digital ecosystems. Managers should adopt outbound practices for knowledge exploitation and technology commercialisation. These dyanmic capabilities are of major importance if firms are to strive long term.

Finally, the work offers a technical and business cycle perspective that can be used by managers. It illustrates the evolutionary path undergone by the Apache Hadoop from community base to enterprise distribution with a service-level agreement. Additionally, the technical and business cycle is divided into the innovation classes (discontinuous, architectural and incremental). This, in turn, enables managers to understand how mature their platforms are and what type of innovation is most suitable for the platform they run.

This study provides insights to firms that are planning to implement the Apache Hadoop to remain competitive in the market. Most importantly, it draws the attention of firms to the importance of knowledge externalisation and technology commercialisation. As more and more firms are shifting to data-driven decision making, how firms manage their innovation classes is of paramount importance. This study sheds light in these types of problems. The findings suggest that, in the evolutionary process of innovation, firms should understand what class of innovation they want to achieve and then operate the right distribution.

8.3. LIMITATIONS AND FUTURE RESEARCH AVENUES

This study follows a qualitative research tradition and adopted a multiple case study design for theory building. Based on that, the first limitation relates to the deficiencies of multiple-case study research. Further researchers should include a quantitative analysis to examine the extent to which the two models apply in other digital ecosystems. Further research should apply a large-scale survey or increase the number of firms and the sectors involved. Considering innovation streaming deals with an evolutionary system – discontinuous, architectural and incremental innovation - a longitudinal study could be advantageous. The fieldwork was carried out in Brazil and the sponsor would only support 10 weeks of data gathering. It would be very challenging to establish such lengthy access to an organisation in Brazil being based in the UK.

This research opens up avenues for further research that are of interest to be explored in the future, contributing to the understanding of outbound OI practices in the context of ambidexterity. Future scholars may explore outbound OI practices in relation to value creation in technology alliances within digital ecosystems. Further studies could also access the impact of such technology alliances in terms of financial and non-financial benefits for firms. Future scholars could also explore further the proposed framework, as it could serve as a basis for further empirical studies, particularly from a quantitative perspective.

BIBLIOGRAPHY

Abernathy, W. A. (1978). *The Productivity Dilemma: Roadblock to Innovation in the Automobile Industry*. Baltimore: Johns Hopkins University Press.

Abernathy, W. A. and Utterback, W. J. (1978). Patterns of Industrial Innovation, *Technology Review*, *80*(7), 40–47.

Abernathy, W. J. and Clark, K. B. (1985). Innovation: Mapping the winds of creative destruction, *Research Policy, 14*(1), 3–22. doi: 10.1016/0048-7333(85)90021-6.

Adam, Y., Ong, C. H. and Pearson, A. W. (1988). Licensing as an Alternative to Foreign Direct Investment: An Empirical Investigation, *Journal of Product Innovation Management*, *5*, 32–49.

Adler, P. S., Goldoftas, B. and Levine, D. I. (1999). Flexibility Versus Efficiency? A Case Study of Model Changeovers in the Toyota Production System, *Organization Science*, *10*, 43-68.

Adomavicius, G. et al. (2008). Making Sense of Technology Trends in the Information Technology Landscape, *MIS Quarterly*, *32*, 779–809.

Adomavicius, G. et al. (2007). Technology Roles and Paths of Influences in an Ecosystem Model of Technology Evolution, *Information Technology and Management, 8*, 185–202.

Agrawal, A. (2005). Engaging the Inventor: Exploring Licensing Strategies for University Inventions and the Role of Latent Knowledge, *Strategic Management Journal, 27*(1), 63–79.

Abbasi, A., Sarker, S., & Chiang, R. H. L. (2016). Big Data Research in Information Systems: Toward an Inclusive Research Agenda. *Journal of the Association for Information Systems*, *17*(2).

Abbott, A. (1990). Conceptions of Time and Events in Social Science Methods: Causal and Narrative Approach. *Historical Methods*, *23*(4), 140–150.

Abbott, A. (1992). What Do Cases Do? Some Notes on Activity in Sociological Analysis. In C. C. Ragin & H. S. Becker (Eds.), *What is a Case? Exploring the Foundations of Social Inquiry*. Cambridge University Press.

Adamides, E., & Karacapilidis, N. (2018). Information Technology for Supporting the Development and Maintenance of OI Capabilities. *Journal of Innovation and Management*.

Adler, P. S., Goldoftas, B., & Levine, D. I. (1999). Flexibility versus Efficiency: A Case Study of Model Changeovers in the Toyota Production System. *Organization Science*, *10*, 43–68.

Adner, R., & Kapoor. (2010). Value Creation in Innovation Ecosystems: How the Structure of Technological Interdependence Affects Firm Performance in New Technology Generations. *Strategic Management Journal*, *31*(1), 306–333.

Afuah, A., & Tucci, C. L. (2012). Crowdsourcing as a Solution to Distant Search. *Academy of Management Review*, *37*(3), 355–375.

Ahlstrom, D. (2010). Innovation Growth: How Business Contributes to Society, *Academy of Management Perspectives, 24*(3), 10–23.

Ahmad Zaidi, M. F. and Othman, S. N. (2015). *Structural Ambidexterity vs. Contextual Ambidexterity: Preliminary Evidence from Malaysia*, 2nd International Conference on Multidisciplinary Innovation for Sustainability and Growth (MISG-2015). Kuala Lumpur, Malaysia.

Ahuja, G. (2000). Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study, *Administrative Science Quarterly, 45*(3), 425–455.

Ahuja, G. and Lampert, C. M. (2001) Entrepreneurship in the Large Corporation: A Longitudinal Study of How Established Firms Create Breakthrough Inventions, *Strategic Management Journal, 22*(6/7), 521–543.

Ajamieh, A. et al. (2016). IT Infrastructure and Competitive Aggressiveness in Explaining and Predicting Performance, *Journal of Business Research, 69*(10), 4667–4674.

Al-Debei, M. M. (2010). *The Design and Engineering of Innovative Mobile Data Services: An Ontological Framework Founded on Business Model Thinking*. Brunel University, West London.

Aldrich, H. (1979). *Organizations and Environments*. Englewood Cliffs, NJ: Prentice Hall.

Alegre, J., Lapiedra, R. and Chiva, R. (2006). A Measurement Scale for Product Innovation Performance, *European Journal of Innovation Management, 9*, 333–346.

Alexy, O., George, G. and Salter, A. (2013). Cui bono? The Selective Revealing of Knowledge and Its Implications for Innovative Activity. 270–291, *Academy of Management Review, 38*, 270–291.

Ali, A. (1994). Pioneering Versus Incremental Innovation: Review and Research Propositions, *Journal of Product Innovation Management, 11*(1), 46–61.

Allarakhia, M. and Walsh, S. (2011). Managing Knowledge Assets under Conditions of Radical Change: The Case of the Pharmaceutical Industry, *Technovation, 31*, 105–117.

Allen, R. C. (1983). Collective Innovation, *Journal of Economic Behavior and Organization*, *4*, 1–24.

Allen, T. J. (1977). *Managing the Flow of Technology*. Cambridge, MA: MIT Press.

Allen, T. J. and Cohen, W. M. (1969). Information Flow in Research and Development Laboratories, *Administrative Science Quarterly, 14*(1), 12–19.

Alexey, O., Henkel, J., & Wallin, M. W. (2013). From Closed to Open: Job Role Changes, Individual Predispositions, and the Adoption of Commercial Open-Source Software Development. *Research Policy*, *42*(8), 1325–1340.

Alnuaimi, T., & George, G. (2016). Appropriability and the Retrieval of Knowledge after Spillovers. *Strategic Management Journal*, *37*(7), 1263–1279.

Aloini, D., Lazzarotti, V., Pellegrini, L., & Zerbino, P. (2020). Inside-Out: The Forgotten side of ICT-Enabled OI. *Measuring Business Excellence*, *24*(1), 24–38.

Ancona, D. G., Goodman, P. S., Lawrence, B. S., & Tushman, M. L. (2001). Time: A New Research Lens. *Academy of Management Review*, *26*, 645–663.

Andersson, G. (2013). Rethinking Regional Innovation. *Systemic Practice and Action Research*, *26*(1), 99–110.

Anderson, P. and Tushman, M. (1990). Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change, *Administrative Science Quarterly, 35*, 604–633.

Andries, P. and Faems, D. (2013). Patenting Activities and Firm Performance: Does Firm Size Matter?, *Journal of Product Innovation Management, 30*(6), 1089–1098.

Andriopoulos, C. and Lewis, M. W. (2009). Exploitation-Exploration Tensions and Organizational Ambidexterity: Managing Paradoxes of Innovation, *Organization Science, 20(*4), 696–717.

Appleton, J. V. and King, L. (2002). Journeying from the Philosophical Contemplation of Constructivism to the Methodological Pragmatics of Health Services Research, *Journal of Advanced Nursing, 40(*6), 641–648.

Appleyard, M. M., & Chesbrough, H. (2017). The Dynamics of Open Strategy: From Adoption to Reversion. *Long Range Planning* , *50*, 310–321.

Archibugi, D., Howells, J. and Michie, J. (1999). I*nnovation Policy in a Global Economy.* Cambridge: Cambridge University Press.

Ardito, L. et al. (2018). The influence of Inbound OI on Ambidexterity Performance: Does it Pay to Source Knowledge from Supply Chain Stakeholders?, *Journal of Business Research*, In Press.

Argyres, N., Mahoney, J. T., & Nickerson, J. (2019). Strategic responses to shocks: Comparative adjustment costs, transaction costs, and opportunity costs. *Strategic Management Journal*, *40*(3), 357–376.

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.

Arora, A., Fosfuri, A. and Gambardella, A. (2001). *Markets for Technology*. Cambridge, MA: MIT Press.

Asheim, B. T., & Gertler, M. (2005). The Geography of Innovation: Regional Innovation Systems. In J. Fagerberg & D. Mowery (Eds.), *The Oxford Handbook of Innovation* (pp. 291–317). Oxford University Press.

Asheim, B. T., Isaksen, A., & Trippl, M. (2019). *Advanced Introduction to Regional Innovation Systems*. Edward Elgar.

Aspara, J. et al. (2011). Strategic Management of Business Model Transformation: Lessons from Nokia, *Management Decision*, 622–647.

Audretsch, D. B. (1995). Innovation, Growth and Survival, *International Journal of Industrial Organization, 13*(4), 441–457.

Avgerou, C. (2002). IT and Organizational Change: An Institutionalist Perspective, *Information Technology & People, 13*(4), 234–262.

Avison, D. E. and Fitzgerald, G. (2003). *ISDevelopment: Methodologies, Techniques and Tools*. Maidenhead, UK: McGraw-Hill.

Baden-Fuller, C. and Volberda, H. (1997). Strategic Renewal in Large Complex Organizations: A Competence- Based View, in *Competence-Based Strategic Management*. Chichester, UK: John Wiley & Sons, 89–110.

Baker, W. E. and Sinkula, J. M. (2007). Does Market Orientation Facilitate Balanced Innovation Programs? An Organizational Learning Perspective, *The Journal of Product innovation Management, 24(*4), 316–334.

Baldwin, C. Y. (2012). Organization Design for Business Ecosystems, *Special Issue on The Future of Organization Design, Journal of Organization Design,* 1(1), 1-4.

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

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

Bamberg, M. (2012). Narrative Analysis. In H. Cooper (Ed.), *APA Handbook of Research Methods in Psychology*. APA Press.

Bantel, K. A. and Jackson, S. E. (1989). Top Management and Innovations in Banking: Does the Competition of the Top Team Make a Difference?, *Strategic Management Journal, 10*, 107–124.

Baregheh, A., Rowley, J. and Sambrook, S. (2009). Towards a Multidisciplinary Definition of Innovation, *Management Decision, 47*(8), 1323–1339.

Barley, S. R. (1986). Technology as an Occasion for Structuring, *Administrative Science Quarterly, 31*, 78–108.

Barnard, C. I. (1938). *The Functions of the Executives*. 30th Anniv. Cambridge, MA: Harvard University Press.

Barney, J. (2002). *Gaining and Sustaining Competitive Advantage*. Upper Saddle River, NJ.: Prentice Hall.

Barras, R. (1990). Interactive Innovation in Financial and Business Services: The Vanguard of the Service Revolution, *Research Policy, 15*, 161–173.

Barras, R. (1986). Towards a Theory of Innovation in Services, *Research Policy, 15*, 161–173.

Basalla, G. (1988). *The Evolution of Technology*. Cambrigde: Cambridge University Press.

Baskerville, R. L. and Myers, M. D. (2002). Information System as a Reference Discipline, *MIS Quarterly, 26*(1), 1–14.

Bassis, N. F., & Armellini, F. (2018). Systems of Innovation and Innovation Ecosystems: A Literature Review in Search of Complementarities. *Journal of Evolutionary Economics*, *28*, 1053–1080.

Baughn, C. C., Denekamp, J. G., Stevens, J. H., & Osborn, R. N. (1997). Protecting intellectual capital in international alliances. *Journal of World Business*, *32*(2), 103–117.

Baum, J., Korn, H. and Kotha, S. (1995). Dominant Designs and Population Dynamics in Telecommunications Services: Founding and Failure of Facsimile Transmission Service Organization, 1965-1992, *Social Science Research*.

Benbasat, I., Goldstein, D. K. and Mead, M. (1987). The Case Research Strategy in Studies of Information Systems, *MIS Quarterly, 11*(3), 369–385.

Benbasat, I. and Zmud, R. W. (2003). The identity crisis within the IS discipline: Defining and communicating the discipline's core properties, *Management ISQuarterly, 27*(2).

Bengtsson, L., & Ryzhkova, N. (2013). Managing a Strategic Source of Innovation: Online Users. *International Journal of Information Management*, *33*(4), 655–662.

Benkler, Y. (2002). Coase's Penguin, Or, Linux and The Nature of the Firm, *The Yale Law Journal, 112*(3).

Benkler, Y. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven and London: Yale University Press.

Benner, M. J. and Tushman, M. L. (2015). Reflections on the 2013 Decade Award —"Exploitation, Exploration, and Process Management: The Productivity Dilemma Revisited" Ten Years Later, *Academy of Management Review, 40*(4), 497–514.

Benner, M. and Tushman, M. (2002). Exploitation, Exploration, and Process Management: The Productivity Dilemma Revisited, *Academy of Management Review, 28*, 238–256.

Berchicci, L. (2013). Towards an Open R&D System: Internal R&D Investment, External Knowledge Acquisition and Innovative Performance. *Research Policy*, *42*, 117–127.

Bergek, A., Berggren, C., Magnusson, T., & Hobday, M. (2013). Document type Article Green Open Access Source type Journal ISSN 00487333 DOI 10.1016/j.respol.2013.02.009 View more Technological discontinuities and the challenge for incumbent firms: Destruction, disruption or creative accumulation? *Research Policy*, *42*(6–7), 1210–1224.

Berghaus, S. and Back, A. (2016). *Stages in Digital Business Transformation: Results of an Empirical Maturity Study,* in The Mediterranean Conference on IS(MCIS). Nicosia, Cyprus.

Berglund, H., & Sandström, C. (2013). Business Model Innovation from an Open Systems Perspective: Structural Challenges and Managerial Solutions. *International Journal of Product Development*, *40*(2), 16–24.

Bessant, J., Lamming, R., Noke, H., & Phillips, W. (2005). Managing innovation beyond the steady state. *Technovation*, *25*(12), 1366–1376.

Bharadwaj, A. S. A. (2000). A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation, *MIS Quarterly, 24*(1), 169–196.

Bianchi, M., Chiesa, V. and Frattini, F. (2011). Selling Technological Knowledge: Managing the Complexities of Technology Transactions, *Research-Technology Management, 54*(9), 18–26.

Bigelow, L., Nickerson, J., & Park, W.-Y. (2019). When and how to shift gears: Dynamic trade-offs between among adjustment costs, opportunity, and transaction costs in response to an innovation shock. *Strategic Management Journal*, *40*(3), 377–407.

Billinger, S. and Jacobides, M. G. (2006). *Changing the Firm's Digital Backbone: How Information Technology shapes the Boundaries of the Firm*, in 19th Bled eConference eValues June 5 - 7. Bled, Slovenia.

Birkenmeier, B. (2003). *External Deployment of Technologies. Technology and Innovation Management on the Move*. Tschirky Verlag Industrielle Organisation.

Birkinshaw, J., Bessant, J., & Delbridge, R. (2007). Finding, Forming, and Performing: Creating Networks for Discontinuous Innovation. *California Management Review*, *49*(3).

Birkinshaw, J. and Gibson, C. (2004). Building Ambidexterity into an Organization, *MIT Sloan Management Review*, June, 47–55.

Birkinshaw, J., Zimmermann, A., & Raisch, S. (2016). How Do Firms Adapt to Discontinuous Change? Bridging the Dynamic Capabilities and Ambidexterity Perspectives. *California Management Review*, *58*(4), 36–58.

Bitran, G. R., Suri, G., & Shiou, L. S. (2007). The Need for Third-Party Coordination in Supply Chain Governance. *MIT Sloan Management Review*, *48*(3).

Björkdahl, J., & Holmén, M. (2013). Editorial: Business Model Innovation – The Challenges Ahead". *International Journal of Product Development*, *18*(3/4), 213–225.

Boddy, C. R. (2016). Sample size for qualitative research. *Qualitative Market Research: An International Journal*, *19*(4), 426–432.

Boehm, B. W. (2002). Get Ready for Agile Methods, With Care, *Computer, 35*, 64-69.

Bogers, M., Chesbrough, H., & Moedas, C. (2018). OI: Research, Practices and Policies. *California Management Review* , *60*(2), 5–16.

Bogers, M., Zobel, A.-K., Esteve Almirall, A. A., Brunswicker, S., Dahlander, L., Frederiksen, L., Gawer, A., Gruber, M., Haefliger, S., Hagedoorn, J., Hilgers, D., Laursen, K., Magnusson, M. G., Majchrzak, A., McCarthy, I. P., Moeslein, K. M., Nambisan, S., Piller, F. T., Radziwon, A., … ter Wal, A. L. J. (2017). The OI Research Landscape: established

Perspectives and Emerging Themes Across Different Levels of Analysis. *Industry and Innovation* , *24*(1), 8–40.

Bogers, M., Afuah, A., & Bastian, B. (2010). Users as Innovators: A Review, Critique, and Future Research Directions. *Journal of Management* , *36*(4), 857–875.

Bogers, M., & West, J. (2012). Managing Distributed Innovation: Strategic Utilization of Open and User Innovation. *Creativity and Innovation Management*, *21*(1), 61–75.

Bogers, M., Sims, J., & West, J. (2019). What Is an Ecosystem? Incorporating 25 Years of Ecosystem Research. *Academy of Management Annual Meeting Proceedings*, *1*.

Boland, R. (1978). The Process and Product of System Design, *Management Science, 28*(9), 887–898.

Boland, R. and Day, W. (1989). The Experience of System Design: A Hermeneutic of Organisational Action, *Scandinavian Journal of Management, 5,* 87–104.

Bolen, J. M. (1964). The Adoption and Diffusion of Ideas in Agriculture, in H., C. J. (Ed.) *Our Changing Rural Society: Perspectives and Trends*. Ames, IA: Iowa State University.

Boschma, R. (2004). Competitiveness of Regions from an Evolutionary Perspective. *Regional Studies*, *38*(9), 1001–1014.

Boudreau, K. J. (2007). *Does Opening a Platform Stimulate Innovation? The Effect on Systemic and Modular Innovations*. No. 4611-05.

Bouncken, R. B., Plüschke, B. D., Pesch, R., & Kraus, S. (2016). Entrepreneurial Orientation in Vertical Alliances: Joint Product Innovation and Learning from Allies. *Review Management Science*, *10*, 381–409.

Bovet, P. D. and Casati, M. (2001). *Understanding the Linux Kernel*. Sebastopol, CA: O'Reilly.

Bozdogan, K. et al. (1998). Architectural Innovation in Product Development through early Supplier Integration, *R&D Management, 28*(3), 163–173.

Braczyk, H. J. and Heidenreich, M. (1998). Regional Governance Structure in a Globalized World, in Braczyk, H. J. et al. (Eds.), *Regional Innovation Systems*. London: University College London Press.

Bravo, M. I. R., Ruiz-Moreno, A. and Lloréns Montes, F. J. (2018). Examining Desorptive Capacity in Supply Chains: The Role of Organizational Ambidexterity, *International Journal of Operations & Production Management, 38*(2), 534–553.

Brenner, M. J. and Tushman, M. (2008). Process Management, Technological Innovation, and organizational Adaption, in Grover, V. and Markus, M. L. (Eds.): *Business Process Transformation*. Armonk, New York: M. E. Sharpe.

Breschi, S. and Malerba, F. (1997). Sectoral Innovation Systems: Technological Regimes, Schumpeterian Dynamics, and Spatial Boundaries, in Edquist, C. (Ed.): *Systems of Innovation: Technologies, Organizations, and Institutions*. London: Pinter, 130–156.

Brettel, M., Oswald, M. and Flatten, T. (2012). Alignment of Market Orientation and Innovation as a Success Factor: a Five-Country duty, *Technology Analysis & Strategic Management, 24*(2), 151–165.

Bromley, D. A. (2004). Technology Policy, *Technology in Society, 26*(2–3), 455–468.

Bröring, S. and Herzog, P. (2008). Organising New Business Development: OI at Degussa, *European Journal of Innovation Management, 11*(3), 330–348.

Brown, G. W., Dermikan, H. and Mitchell, M. (2006). *Towards the Service-Oriented Enterprise Vision: Bridging Industry and Academics*, in Panel Presentation American Conference on Information System. Acapulco, Mexico.

Brown, J. S. and Duguid, P. (2001). Knowledge and Organization: A Social-Practice Perspective, *Organization Science, 12*, 198–213.

Brown, J. S. and Hagel III, J. (2006). Creation Nets: Getting the Most from OI, *McKinsey Quarterly, 2*, 40–51.

Bruner, J. (1986). *Actual Minds, Possible Worlds*. Harvard University Press.

Bruner, J. (1990). *Acts of Meaning* . Harvard University Press.

Buckley, P. J. and Prashantham, S. (2016). Global Interfirm Networks: The Division of Entrepreneurial Labor between MNEs and SMEs, *Academy of Management Journal, 30*(1), 40–58.

Bugshan, H. (2015). OI Using Web 2.0 Technologies. *Journal of Enterprise Information Management*, *28*(4), 595–607.

Burgelman, R. A. (1991). Intraorganizational Ecology of Strategy Making and Organizational Adaptation: Theory and Field Research. *Organization Science*, *2*, 239–262.

Burkhart, T. et al. (2011). *Analyzing the Business Model Concept – A Comprehensive Classification of Literature*, in Proceedings of the International Conference on Information Systems, ICIS 2011. Shanghai, China.

Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Bustina, O. F., Venderll-Herrero, F. and Gomes, E. (2019). Unpacking the Effect of Strategic Ambidexterity on Performance: A Cross- Country Comparison of MMNEs Developing Product-Service Innovation, *International Business Review*, in Press.

Caputo, M., Lamberti, E., Cammarano, A., & Michelino, F. (2016). Exploring the Impact of OI on Firm Performances. *Management Decision*, *54*(7), 1788–1812.

Cainelli, G., Evangelista, R. and Savona, M. (2004). The Impact of Innovation on EconomicPerformance in Services, *The Service Industries Journal, 24*(1), 116–130.

Calantone, A., Roger, J. and Di Benedetto, C. (1988). An Integrative Model of the New Product Development Process: An Empirical Validation, *Journal of Product Innovation Management, 5*, 201–215.

Caldas Vianna, B. (2021). Knowledge art: Artwork and invention. *Artech 2021-Proceedings of the 10th International Conference on Digital and Interactive Arts. Association for Computing Machinery*.

Camagni, R. (1991). *Innovation Networks: Spatial Perspectives*. London: Belhaven.

Campbell, D. T. (1969). Variation and selective reten- tion in socio-cultural evolution, *General Systems, 14*, 69–85.

Cantwell, J. (2004). An historical change in the nature of corporate technological diversification, in Cantwell, J., Garmbardella, A., and Granstrand, O. (Eds.): *The Economies and Management of Technological Diversification*. London, UK: Routledge, 265–296.

Cao, Q., Gedajlovic, E. and Zhang, H. (2009). Unpacking Organizational Ambidexterity: Dimensions, Contingencies and Synergistic Effects, *Organization Science, 20*(4), 781–796.

Capra, E., Francalanci, C. and Merlo, F. (2008). *An Empirical Study on the Relationship between Software Design Quality, Development Effort and Governance in Open Source Projects*, IEEE Transactions on Software Engineering, 765–782.

Carayannis, E. G. and Rakhmatullin, R. (2014). The Quadruple/Quintuple Innovation Helixes and Smart Specialisation Strategies for Sustainable and Inclusive Growth in Europe and Beyond, *Journal of the Knowledge Economy, 5*(2), 212–239.

Carlsson, B. and Stankiewicz, R. (1991). On the Nature, Function and Composition of Technological Systems, *J Evol Econ, 1*, 93–118.

Carter, C. F. and Williams, B. R. (1959). The Characteristics of Technically Progressive Firms, *Journal of Industrial Economics, March,* 87–104.

Carver, B. W. (2005). Share and Share Alike: Understanding and Enforcing Open Source and Free Software Licenses, *Berkeley Technology Law Journal, 20*, 442-481.

Cascio, W. F., & Aguinis, H. (2005). *Applied Psychology in Human Resource Management*. Pearson Prentice Hall.

Cassiman, B. and Valentini, G. (2016). OI: Are Inbound and Outbound Knowledge Flows Really Complementary?, *Strategic Management Journal, 37*(6), 1034–1046.

Cavalcante, S., Kesting, P. and Ulhøi, J. (2011). Business Model Dynamics and Innovation: Re-establishing the Missing Linkages, *Management Decision, 49*, 1327–1342.

Ceccagnoli, M. et al. (2012). Cocreation of Value in a Platform Ecosystem: The Case of Entreprise Software, *MIS Quarterly, 36*, 263–290.

Chae, B. (2019). A General Framework for Studying the evolution of the Digital Innovation Ecosystem: The Case of Big Data. *International Journal of Information Management*, *45*, 83–94.

Chae, B. (2014). A Complexity Theory Approach to IT-Enabled Services (IESs) and Service Innovation: Business Analytics as an Illustration of IES. *Decision Support Systems*, *57*, 1–10.

Chae, S., Yan, T., & Yang, Y. (2020). Supplier innovation value from a buyer–supplier structural equivalence view: Evidence from the PACE awards in the automotive industry. *Journal of Operations Management*, *66*(7–8), 820–838.

Chandler, A. (1966). *Strategy and Structure: Chapters in the History of the Industrial Enterprise*. Garden City, N.Y.: Anchor.

Chandler, A. D. (1977). *The Visible Hand: The Managerial Revolution in American Business*. Cambrigde, MA: Harvard University Press.

Chang, Y.-C. and Chen, M.-H. (2004). Comparing Approaches to Systems of Innovation: The Knowledge Perspective, *Technology in Society, 26*, 17–37.

Chatman, S. (1978). *Story and Discourse: Narrative Structure in Fiction and Film*. Cornell University Press.

Cheah, S. L.-Y., & Ho, Y.-P. (2021). Commercialization Performance of Outbound OI Projects in Public Research Organizations: The Roles of Innovation Potential and Organizational Capabilities. *Industrial Marketing Management*, *94*, 229–241.

Chen, C. C. J. and Shiu, E. C. (2019). Establishing a Typology of OI Strategies and their Differential Impacts on Innovation Success in an Asia-Pacific Developed Economy, *Asia Pacific Journal of Management*.

Chen, J., Chen, Y., & Vanhaverbeke, W. (2011). The Influence of Scope, depth, and Orientation of External Technology Sources on the Innovative Performance of Chinese Firms. *Technovation*, *31*, 362–373.

Cheng, C. . C. J., Yang, C. and Sheu, C. (2016). Effects of OI and Knowledge-Based Dynamic Capabilities on Radical Innovation: An Empirical Study, *Journal of Engineering and Technology Management, 41*, 79–91.

Chen, J., Chen, Y., & Vanhaverbeke, W. (2011). The Influence of Scope, depth, and Orientation of External Technology Sources on the Innovative Performance of Chinese Firms. *Technovation*, *31*, 362–373.

Cheng, C. and Huizingh, K. R. E. (2010). OI to Increase Innovation Performance: Evidence rom a Large Survey, in Huizingh, K. R. E. et al. (Eds.): *Proceedings of the XXI ISPIM International Conference*. Bilbao, Spain.

Chesbrough, H. and Spohrer, J., H. (2006). A research manifesto for services science, *Communication of the ACM, 49*(7).

Chesbrough, H. and Crowther, A. K. (2006). Beyond High Tech: Early Adopters of OI in other Industries, *R&D Management, 36*, 229–236.

Chesbrough, H. (2007). Business Model Innovation: It's not just about Technology Anymore, *Strategy & Leadership, 35*(6), 12–17.

Chesbrough, H. (2003). The Governance and Performance of Xerox's Technology Spin-off Companies, *Research Policy, 32*, 403–421.

Chesbrough, H. (2006). Innovation, OI: A New Paradigm for Understanding Industrial, in Chesbrough, H., Vanhaverbeke, W., and West, J. (Eds.): *OI: Researching a New Paradigm*. Oxford, United Kingdom: Oxford University Press.

Chesbrough, H. (2003). The Logic of OI: Managing Intellectual Property, *California Management Review, 45*(3), 33–58.

Chesbrough, H. (2003). *OI: The New Imperative for Creating and Profiting from Technology*. Boston, Massachusetts: Harvard Business School Press.

Chesbrough, H. (2006). A Research Manifesto for Services Science, *Communications of the ACM, 49*(7), 35–40.

Chesbrough, H. (2007). Business Model Innovation: It's not Just About Technology Anymore, *Strategy & Leadership, 35*(6), 2007.

Chesbrough, H. (2004). Managing OI, *Research Technology Management, 47*(1), 23–26.

Chesbrough, H. (2010). Business Model Innovation: Opportunities and Barriers, *Long Range Planning, 43,* 354–363.

Chesbrough, H. (2006). *OI: The New Imperative for Creating and Profiting from Technology.* Boston, Massachusetts: Harvard Business School Press.

Chesbrough, H. W. (2003). The Logic of OI: Managing Intellectual Property, *California Management Review, 45*(3), 33–58.

Chesbrough, H. W. (2007). Why Companies Should Have Open Business Models, *MIT Sloan Management Review, 48*(2), 22–28.

Chesbrough, H. W. (2003). The Era of OI, *MIT Sloan Management Review Management Review, 44*(3), 35–41.

Chesbrough, H., and Crowther, A. K. (2006). Beyond High Tech: Early Adopters of OI in Other Industries, *R&D Management, 36*(3), 229–236.

Chesbrough, H., Garman, A. (2009). How OI Can Help You Cope in Lean Times, *Harvard Business Review, 87*, 68–76.

Chesbrough, H. and Appleyard, M. (2007). OI and Strategy, *California Management Review, 50*(1), 57–76.

Chesbrough, H. and Bogers, M. (2014). Explicating OI: Clarifying an Emerging Paradigm for Understanding Innovation, in Chesbrough, H., Vanhaverbeke, W., and West, J. (Eds.): *New Frontiers in OI*. Oxford Scholarship Online.

Chesbrough, H. and Brunswicker, S. (2014). A Fad or a Phenomenon?: The Adoption of OI Practices in Large Firms, *Research-Technology Management, 57(*2), 16–25.

Chesbrough, H., Lettl, C., & Ritter, T. (2018). Value Creation and Value Capture in OI. *Journal of Product Innovation Management*, *35*, 930–938.

Chesbrough, H., Lettl, C., & Ritter, T. (2018). Value Creation and Value Capture in OI. *Product Development & Management Association*, *35*(6), 930–938.

Chesbrough, H. and Rosenbloom, R. (2002). The Role of the Business Model in Capturing Value from Innovation: Evidence from Xerox Corporation's Technology Spin-off Companies, *Industrial and Corporate Change, 11*(3), 529–555.

Chesbrough, H. and Schwartz, K. (2007). Innovating Business Models with Co-Development Partnerships, *Research-Technology Management, 500*(1), 55–59.

Chi, M. et al. (2017). The Influence of Inter-Firm IT Governance Strategies on Relational Performance: The Moderation Effect of Information Technology Ambidexterity, *International Journal of Information Management, 37(*2), 43–53.

Chiaromonte, F. (2006). OI Through Alliances and Partnership: Theory and Practice, *International Journal of Technology Management, 33*(2–3), 111–114.

Chiaroni, D., Chiesa, V. and Frattini, F. (2009). Investigating the Adoption of OI in the Bio-pharmaceutical Industry: A Framework and an Empirical Analysis, *European Journal of Innovation Management, 12*(3), 285–305.

Chou, C., Yang, K.-P. and Chiu, Y.-J. (2016). Coupled OI and Innovation Performance Outcomes: Roles of Absorptive Capacity, *Corporate Management Review, 36*(1), 37–68.

Christensen, C. M. (1997). *The innovator's dilemma*. Boston, MA: Harvard Business Review Press.

Christensen, C. M. and Bower, J. L. (1996). Customer Power, Strategic Investment, and the Failure of Leading Firms, *Strategic Management Journal, 17*(3), 197–218.

Christensen, C. M. and Raynor, M. E. (2003). *The Innovator's solution: Creating and sustaining successful growth*. Harvard Business School Press.

Christensen, C. M. and Rosenbloom, R. S. (1995). Explaining the attacker's advantage: Technological paradigms, organizational dynamics, and the value network, *Research Police, 24(*2), 233–257.

Christensen, J. F., Olesen, M. H. and Kjær, J. S. (2005). The Industrial Dynamics of OI: Evidence from the Transformation of Consumer Electronics, *Research Policy, 34*(10), 1533–1549.

Chua, W. F. (1986). Radical Developments in Accounting Thought, *Accounting Review, 61,* 601–632.

Chung, S. (2002). Building a National Innovation System through Regional Innovation Systems. *Technovation*, *22*(8), 485–491.

Clark, C. (1957). *Conditions of Economic Progress*. New York: Macmillan.

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

Clark, K. B. and Fujimoto, T. (1989). *Product Development Performance*. Cambridge, MA: Harvard Business School Press.

Cleary, M., Horsfall, J. and Hayter, M. (2014). Data Collection and Sampling in Qualitative Research : Does Size Matter?, *Journal of Advanced Nursing, 70*(3), 473–475.

CMMI Product Team (2006). *Capability Maturity Model Integration for Development (CMMI-DEV)* [online]. Available at: http://www.sei.cmu .edu/cmmi/tools/dev/index.cfm

Cohen, S. S. and Zysman, J. (1988). *Manufacturing Matters: The Myth of the Post-Industrial Economy*. New York: Basic.

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

Conlon, C., Timonen, V. and Elliot-O'Dare, C. (2020). Confused About Theoretical Sampling? Engaging Theoretical Sampling in Diverse Grounded Theory Studies, *Qualitative Health Research, 30*(6), 947–959.

Connell, J., & Voola, R. (2013). Knowledge Integration and Competitiveness: A Longitudinal Study of an Industry Cluster. *Journal of Knowledge Management*, *17*(2), 208–225.

Connelly, F. M., & landinin, D. J. (1990). Stories of experience and narrative inquiry. *Educational Researcher*, *19*(5), 2–14.

Conner, D. R. (1995). Obtaining Strategic Advantage from Being Imitated: When Can Encouraging "Clones" Pay?, *Management Science, 41*, 209–225.

Connor, T. (1999). Customer-Led and Market-Oriented: A Matter of Balance, *Strategic Management Journal Management Journal, 20*(12), 1157–1163.

Constant, E. W. (1987). The social locus of techno- logical practice: Community, system in organization, in Bijker, W. E., Hughes, P. T., and Pinch, T. J. (Eds.): *The Social Construction of Technological Systems*. Cambridge, MA: MIT Press.

Constant, E. W. (1980). *The Origins of the Turbojet Revolution*. Baltimore: Johns Hopkins University Press.

Converse, P. D. (1930). *The Elements of Marketing*. New York: Prentice-Hall.

Cooke, P. (1992). Regional innovation systems: Competitive Regulation in the New Europe. *Geoforum; Journal of Physical, Human, and Regional Geosciences*, *23*(3), 365–382.

Cooke, P. (2005). Regionally Asymmetric Knowledge Capabilities and OI Exploring "Globalisation 2" - A New Model of Industry Organisation, *Research Policy, 34*(8), 1128–1149.

Cooke, P., Uranga, M. and Etexbarria, G. (1997). Regional Innovation Systems: Institutional and Organizational Dimension, *Research Policy, 26*, 475–491.

Cowan, R. and Jonard, N. (2004). Network Structure and the Diffusion of Knowledge, *Journal of Economic Dynamics and Control, 28*(8), 1557–1575.

Craighead, C. W., Hult, G. T. M., & Ketchen, D. J. (2009). The effects of innovation–cost strategy, knowledge, and action in the supply chain on firm performance. *Journal of Operations Management*, *27*(5), 405–421.

Crawford, C. M. (1994). *Significant Issues for the Future of Product Innovation*. Boston: Irwin, Burr Ridge.

Crawford, C. M. (1980). Defining the Charter for Product Innovation, *Sloan Management Review, 22(*1).

Creswell, J. W. (2013). *Qualitative Inquiry and Research Design: Choosing Five Approaches*. Thousand Oaks, CA: Sage Publications, Inc.

Crosswhite, J. (1996). *The Rhetoric of Reason*. University of Winsconsin Press.

Crotty, M. (1998). *The Foundations of Social Research: Meaning and Perspective in the Research Process*. London: Sage Publications.

Czarniawska, B. (2004). *Narratives in Social Science Research. Introducing Qualitative* . Sage Publications.

Czarniawska, B., & Joerges, B. (1995). Winds of Organizational Change: How Ideas Translate into Objects and Actions. In S. B. Bacharach, P. Gagliardi, & B. MUndell (Eds.), *Research in the Sociology of Organizations* (Vol. 13). JAI Press.

Dabrowska, J., Lopez-Vega, H., & Ritala, P. (2019). Waking the Sleeping Beauty: Swarovski's OI Journey. *R&D Management*, *49*(5), 775–788.

D'Adderio, L. (2014). The Replication Dilemma Unravelled: How Organizations Enact Multiple Goals in Routine Transfer, *Organization Science, 25*, 1325–1350.

Dalcher, D., & Drevin, L. (2004). Learning from ISFailures by Using Narrative and Ante-Narrative Methods : Research Article. *South African Computer Journal*, *33*.

D'Aveni, R. A. (1994). *Hypercompetition: Managing the Dynamics of Strategic Maneuvering.* New York: The Free Press.

Davidson, E. J. (1997). Examining Project History Narratives: An Analytic Approach. *International Conference on ISand Qualitative Research*, 123–148.

Daft, R. L. (1978). Dual-Core Model of Organizational Innovation, *Academy of Management Journal, 21*(2), 193–210.

Dahlander, L. and Gann, D. M. (2010). How Open is Innovation?, *Research Policy, 39*(6), 699–709.

Dahlander, L. and Magnusson, M. (2008). How do Firms Make Use of Open Source Communities?, *Long Range Planning, 41*(6), 629–649.

Dahlander, L. and Wallin, M. W. (2006). A Man on the Inside: Unlocking Communities as Complementary Assets, *Research Policy, 35*, 1243–1259.

Dahmén, E. (1988). Developing Blocks' in Industrial Economics, *Scandinavian Economic History Review, 1,* 3–14.

Dai, Y. et al. (2019). Learning from Referents Across Geographic Distance and Institutional Contexts: The Perceived Speed of ISO 9000 Certification in Chinese High Technology New Ventures, *Journal of Small Business Management, 57*(2), 2019.

Damanpour, F. (1991). Organizational Innovation: A Meta-Analysis of Effects of Determinants and Moderators, *Journal of Management, 34*(3), 555–590.

Damanpour, F. (1996). Organizational Complexity and Innovation: Developing and Testing Multiple Contingency Models, *Manage Science, 42*(5), 693–716.

Damanpour, F. and Evan, W. M. (1984). Organizational Innovation and Performance: The Problem of "Organizational Lag", *Administrative Science Quarterly, 29*(3), 392–402.

Danneels, E. (2002). The Dynamics of Product Innovation and Firm Competences, *Strategic Management Journal, 23*(12), 1095–1121.

Darke, P., Shanks, G. and Broadbent, M. (1998). Successfully Completing Case Study Research: Combining Rigour, Relevance and Pragmatism, *Information System Journal, 8*, 272–289.

David, P. (1987). Some New Standards for the Economics of Standardization in the Information Age, in *Economic Policy and Technological Performance*. Cambridge: Cambridge University Press, 206–239.

David, P. (1985). Clio and the Economics QWERTY, *Economic History, 75*, 332–337.

Davies, A., Brady, T. and Hobday, M. (2007). Organizing for Solutions: Systems Seller vs Systems Integrator, *Industrial Marketing Management, 36*(2), 183–193.

Davis, J. L. and Harrison, S. S. (2001) Edison in the Boardroom: How Leading Companies Realize Value from Their Intellectual Assets. New York, NY: John Wiley & Sons.

de Araújo Burcharth, A. L., Knudsen, M. P. and Søndergaard, H. A. (2014). Neither Invented nor Shared Here: The Impact and Management of Attitudes for the Adoption of OI Practices, *Technovation, 34*, 149–161.

de Camargo Fiorini, P., Pais Seles, B. M. R., Chiappetta Jabbour, C. J., Barbeiro Mariano, E. and Lopes de Souza Jabbour, A. B. (2018). Management Theory and Big Data Literature: From a Review to a Research Agenda. *International Journal of Information Management*, *43*, 112–129.

de Jong, J. P. J. and Marsili, O. (2006). The Fruit Flies of Innovations: A Taxonomy of Innovative Small Firms. *Research Policy*, *35*(2), 213–229.

De la Mothe, J. and Paquet, G. (1996). *Evolutionary Economics and the new International Political Economy*. London: Pinter.

Deci, E. L. and Ryan, R. M. (2008). Self-Determination Theory: A Macro Theory of Human Motivation, Development, and Health, *Canadian Psychology/Psychologie Canadienne, 49*(3), 182–185.

Del Guidice, M., Carayannis, E. G., Palacios-Márques, D., Soto-Acosta, P., & Meissner, D. (2018). The Human Dimension of Open Innovation. *Management Decision*, *56*(6), 1159–1166.

Dell' Anno, D. and del Giudice, M. (2015). Absorptive and Desorptive Capacity of Actors within University-Industry Relations: Does Technology Transfer Matter?, *Journal of Innovation and Entrepreneurship, 4*(13), 1–20.

Deloitte (2020). *Deloitte Public Sector Analytics: Informed Decision-Making* [online]. Available at: https://www2.deloitte.com/uk/en/pages/public-sector/solutions/public-sector-analytics.html

Demil, B. and Lecocq, X. (2010). Business Model Evolution: In Search of Dynamic Consistency, *Long Range Planning, 43*, 227–246.

Demirkan, H. et al. (2008). Service-Oriented Technology and Management: Perspectives on Research and Practice for the Coming Decade, *Electronic Commerce Research and Applications, 7*(4), 356–376.

Den Hertog, P. (2000). Knowledge-Intensive Business Services as Co-Producers of Innovation, *International Journal of Innovation Management, 4*, 491–528.

De Reyck, B. Grushka-Cockayne, Y., Lockett, M. Calderini, S. R., Moura, M. and Sloper, A. (2005). The Impact of Project Portfolio Management on Information Technology Projects, *International Journal of Project Management, 23*, 524-537.

Derbyshire, J. (2014). The Impact of Ambidexterity on Enterprise Performance: Evidence from 15 Countries and 14 Sectors. *Technovation*, *34*, 574–581.

Dermikan, H. and Goul, M. (2006). Towards the Service Oriented Enterprise Vision: Bridging Industry and Academics, *Communication of the ACM, 18.*

Dewar, R. D. and Dutton, J. E. (1986). The adoption of radical and incremental innovations: An empirical analysis, *Management Science, 32*(11), 1422–1433.

Díaz Andrade, A. (2009). Interpretive Research Aiming at Theory Building: Adopting and Adapting the Case Study Design, *The Qualitative Report, 14*(1), 42–60.

di Gangi, P. M., & Wasko, M. (2009). Steal my Idea! Organizational Adoption of User Innovations from a User Innovation Community: A Case Study of Dell IdeaStorm. *Decision Support Systems*, *48*(1), 303–312.

Dittrich, K. and Duysters, G. (2007). Networking as a Means to Strategy Change: The Case of OI in Mobile Telephony, *Journal of Product Innovation Management, 24*(5), 510–521.

Dodgson, M., Gann, D. and Salter, A. (2006). The Role of Technology in the Shift Towards OI: The Case of Procter & Gamble, *R&D Management, 36*(3), 333–343.

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

Dosi, G., & Nelson, R. R. (2010). Technical change and industrial dynamics as evolutionary processes. In B. H. Hall & N. Rosenberg (Eds.), *Handbook of the Economics of Innovation* (1st ed.). North-Holland.

Dougherty, D. and Dunne, D. D. (2011). Organizing Ecologies of Complex Innovation, *Organization Science, 22*(5), 1214–1223.

Dougherty, D. and Heller, T. (1994). The Illegitimacy of Successful Product Innovation in Established Firms, *Organization Science, 5*(2), 79–201.

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

Drechsler, W. and Natter, M. (2012). Understanding a Firm's Openness Decisions in Innovation, *Journal of Business Research, 65*, 438–445.

Duchesneau, T. D., Cohn, S. F. and Dutton, J. E. (1979). *A Study of Innovation in Manufacturing: Determinants, Processes, and Methodological Issues*. Orono, ME: Social Science Research Institute, University of Maine.

Duncan, R. (1976). The ambidextrous Organization: Designing dual Structures for Innovation, in Kilman, R., Pondy, L., and Slevin, D. (Eds.): *The Management of Organizational*. New York: North Holland.

Dutton, J. and Thomas, A. (1985). Relating Technological Change and Learning by Doing', in Rosenbloom, R. D. (Ed.): *Research on Technological Innovation, Management, and Policy. Greenwich*, CT: JAI Press, 187–224.

Dyer, J. (1996). Does Governance Matter? Keiretsu Alliances and Asset Specificity as Sources of Japanese Competitive Advantage, *Organization Science, 7*, 649–666.

Dyer, J. H. and Singh, H. (1998). The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage, *Academy of Management Review, 4,* 660–679.

Dworkin, S. (2012). Sample size policy for qualitative studies using in-depth interviews. *Archives of Sexual Behavior*, *41*(6), 1319–1320.

Easterby-Smith, M., Lyles, M., & Tsang, E. (2008). Inter-organizational Knowledge Transfer: Current Themes and Future Prospect. *Journal of Management Studies*, *45*, 677–690.

Edmondson, A. C. and McManus, S. E. (2007). Methodological Fit in Management Field Research, *Academy of Management Review, 32*(4), 1155–1179.

Edquist, C. (1997). *Systems of Innovation: Technologies, Institutions and Organizations.* London: Pinter.

Edquist, C. and Hommen, L. (1999). Systems of Innovation: Theory and Policy for the Demand Side, *Technology In Society, 21*, 63–79.

Edquist, C. (2001). The Systems of Innovation Approach and Innovation Policy: An Account of the State of the Art. *DRUID Conference*, 12–15.

Eggers, J. P., & Francis Park, K. (2018). Incumbent Adaptation to Technological Change: the Past, Present, and Future of Research on Heterogeneous Incumbent Response. *Academy of Management Annals*, *12*, 257–389.

Ehrhardt, M. (2004). Network Effects, Standardisation and Competitive Strategy: How Companies Influence the Emergence of Dominant Designs, *International Journal of Technology Management, 27*, 272–294.

Eisenhardt, K. M. (1989). Building Theories from Case-Study Research, *Academy of Management Review, 4*(4), 532–550.

Eisenhardt, K. M. and Martin, J. A. (2000). Dynamic Capabilities: What are They?, *Strategic Management Journal, 21*, 1105–1121.

Eisenmann, T. R., Parker, G. and van Alstyne, M. W. (2009). Opening Platforms: How, When and Why?, in *Platforms, Markets and Innovation. Northampton*, MA: Edward Edgar.

Elliot, J. (2005). *Using Narrative in Social Research. Qualitative and Quantitative Approaches*. Sage Publications.

El Sawy, O. A. et al. (2010). Seeking the Configurations of Digital Ecodynamics: It Takes Three to Tango, *ISResearch, 21*(4), 835–848.

Elmquist, M., Fredberg, T. and Ollila, S. (2009). Exploring the Field of OI, *European Journal of Innovation Management Management, 12(*3), 326–345.

Enkel, E. and Gassmann, O. (2010). Creative Imitation: Exploring the Case of Cross-Industry Innovation, *R&D Management, 40*(3), 256–270.

Enkel, E., Gassmann, O. and Chesbrough, H. (2009). Open R&D and OI: Exploring the Phenomenon, *R&D Management, 39*(4), 311–316.

Enkel, E., Heil, S., Hengstler, M. and Wirth, H. (2017). Exploratory and exploitative innovation: To what extent do the dimensions of individual level absorptive capacity contribute? *Technovation*, *60–61*, 29–38.

Enkel, E., Kausch, C. and Gassmann, O. (2005). Managing the Risk of Customer Integration, *Journal European Management, 23*(2), 203–213.

Enkel, E. and Lenz, A. (2009). *OI Metrics Systems*, in R&D Management Conference. Vienna, Austria.

Escher, J.-P. (2005). *Technology Marketing in Technology-Based Enterprises: The Process and Organization Structure of External Technology Deployment*. Eidgenössische Technische Hochschule, Zürich.

Ethiraj, S. K., Levinthal, D., & Roy, R. R. (2008). The Dual Role of Modularity: Innovation and Imitation. *Management Science*, *54*(5), 939–955.

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.

Evangelista, R. (2000). Sectoral patterns of technological Change in Services. *Economics of Innovation and New Technology*, *9*(3), 183–222.

Evan, W. M. (1966). Organizational Lag, *Human Organization, 25*(1), 51–53.

Everett, M. R. (1962). *Diffusion of innovations*. New York: Macmillan Co.

Faems, D., B., Van Looy, B. and Debackere, K. (2009). Interorganizational Collaboration and Innovation: Toward a Portfolio Approach, *Journal of Product Innovation Management, 22*(3), 238–250.

Fagerberg, J. (2004). Innovation: A Guide to the Literature, in Fagerberg, J., Mowery, D., and Nelson, R. (Eds.) *The Oxford Handbook of Innovation*. Oxford: Oxford University Press.

Fartes dos Santos e Silva, D., Bomtempo, J. V., & Chaves Alves, F. (2019). Innovation Opportunities in the Brazilian Sugar-Energy Sector. *Journal of Cleaner Production*, *218*, 871–879.

Faulkner, P. and Runde, J. (2009). On the Identity of Technological Objects and Users Innovation in Functions, *Academy of Management Review, 34*, 442–462.

Feeny, D. F. and Wilcocks, L. P. (1998). Core IS Capabilities for Exploiting Information Technology, *Sloan Management Review, 39*, 9–21.

Feller, J., Finnegan, P., & Nilsson, O. (2011). OI and Public Administration: Transformational Typologies and Business Model Impacts. *European Journal of Information Systems*, *20*(3), 358–374.

Fernandes, C., Farinha, L., Ferreira, J. J., Asheim, B., & Rutten, R. (2020). Regional Innovation Systems: What Can We Learn from 25 Years of Scientific Achievements? *Regional Studies*, *55*(3), 377–389.

Fernández, E., & Valle, S. (2019). Battle for dominant design: A decision-making model. *European Research on Management and Business Economics*, *25*, 72–78.

Ferraris, A. and Santoro, G. (2014). Come dovrebbero Essere Sviluppati i Progetti Di Social Innovation nelle Smart City? Un’Analisi Comparativa, *Impresa Progetto-Electronic Journal of Management, 4*, 1–15.

Ferrary, M. (2011). Specialized Organizations and Ambidextrous Clusters in the OI Paradigm, *European Management Journal, 29*(3), 181–192.

Fetterhoff, T. J. and Voelkel, D. (2006). Managing OI in Biotechnolgy, *Research-Technology Management, 49(*3), 14–18.

Fey, C. and Birkinshaw, J. (2005). External Sources of Knowledge, Governance Mode and R&D Performance, *Journal of Management, 31*(4), 597–621.

Filiou, D. (2020). A New Perspective on OI: Established and New Technology Firms in UK Bio-Pharmaceuticals. *R&D Management*, *51*(1), 73–86.

Filser, M., Kraus, S., Breier, M., Nenova, I., & Puumalainen, K. (2020). Business Model Innovation: Identifying Foundations and Trajectories. *Business Strategy and the Environment*, *30*(2), 891–907.

Finnegan, R. (1992). Oral Traditions and the Verbal Arts: A Guide to Research Practices. In C. Griffin, A. Tindley, & T. Williamson (Eds.), *Rural History Economy* (Vol. 3). Cambridge University Press.

Fitzgerald, B. (2006). The transformation of open source software, MIS Quarterly, 3(3), 587–598.

Fitzgerald, B. and Howcroft, D. (1998). Towards Dissolution of the IS Research Debate: From Polarisation to Polarity 1, *Journal of Information Technology, 13*(4), 1–23.

Fleming, L. (2001). Recombinant Uncertainty in Technological Search, *Manage Science, 47*(1), 117–132.

Fleming, L. and Waguespack, D. M. (2007). Brokerage, Boundary Spanning, and Leadership in OI Communities, *Organization Science, 18*(2), 165–180.

Florén, H. and Frishammar, J. (2012). From Preliminary Ideas to Corroborated Product Definitions: Managing the Front End of New Product Development, *California Management Review, 54*(4), 20–43.

Fonstad, N. O. and Robertson, D. (2006). Transforming a Company, Project by Project, *MIS Quarterly Executive, 5*, 1–14.

Ford, D. (1985). The management and marketing of technolog. In R. Lamb & P. Shrivastava (Eds.), *Advances in Strategic Management*. JAI Press.

Fosfuri, A. (2006). The Licensing Dilemma: Understanding the Determinants of the Rate of Technology Licensing, *Strategic Management Journal, 27*, 1141–1158.

Foss, N. J., & Saebi, T. (2017). Fifteen Years of Research on Business Model Innovation: How Far Have We Come, and Where Should We Go? *Journal of Management*, *43*(1), 200–227.

Francis, D. and Bessant, J. (2005). Targeting Innovation and Implications for Capability Development, *Technovation, 25*(3), 171–183.

Franck, E. and Jungwirth, C. (1999). Die Governance von Open-Source-Projekten, *Ökonomische Analyse von Governance-Strukturen, 5*, 1–21.

Franke, N. and Shah, S. (2003). How Communities Support Innovative Activities: An Exploration of Assistance and Sharing Among End-Users, *Research Policy, 32*(1), 157–178.

Fredberg, T., Elmquist, M. and Ollila, S. (2008). *Managing OI – Present Findings and Future Directions.*

Freeman, C. (1982). *The Economics of Industrial Innovation*. London : Frances Pinter.

Freeman, C. (1987). *Technology Policy and Economic Performance: Lesson From Japan.* London: Frances Pinter.

Frey, K., Lüthje, C., & Haag, S. (2011). Whom Should Firms Attract to OI Platforms? The Role of Knowledge Diversity and Motivation. *Long Range Planning*, *44*(5–6), 397–420.

Frisk, R. P., Brown, S. W. and Bitner, M. J. (1993). Tracking the Evolution of the Services Marketing Literature, *Journal of Retailing, 69*(1), 61–103.

Frøslev, C. J. et al. (2005). The Industrial Dynamics of OI - Evidence from the Transformation of Consumer Electronics, *Research Policy, 34*(10), 1533–1549.

Fruin, W. M. (1996). The Japanese Enterprise System: Competitive Strategies and Cooperative Strategies, *Asia Pacific Journal of Human Resources, 34(*2), 132–133.

Gadrey, J. and Gallouj, F. (2002). *Productivity, Innovation and Knowledge in Services: New Economic & Socio-Economic Approaches*. Cheltenham, UK: Edward Edgar.

Gambardella, A., Giuri, P. and Luzzi, A. (2007). The Market for Patents in Europe, *Research Policy, 36*(8), 1163–1183.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and Analytics. *International Journal of Information Management*, *35*(2), 137–144.

Gann, D. M. (2005). OI: The New Imperative for Creating and Profiting from Technology, *Research Policy, 34*(1), 122–123.

Gans, J., Hsu, D. and Stern, S. (2001). *When Does Startup Innovation Spur the Gale of Creative Destruction*.

Garud, R., Jain, S. and Tuertscher, P. (2008). Incomplete by Design and Designing for Incompleteness, *Organization Science, 29*, 351–371.

Garud, R., Tuertscher, P., & van de Ven, A. H. (2013). Perspectives on Innovation Processes. *The Academy of Management Annals*, *7*(1), 775–819.

Gaskin, J., Berente, N., Lyytinen, K., & Rose, G. (2017). Innovation Among Different Classes of Software Development Organizations. *Information Systems Journal*, *28*, 849–878.

Gassmann, O. (2006). Opening up the Innovation Process: Towards an Agenda, *R&D Management, 36*(3), 223–228.

Gassmann, O. and Enkel, E. (2004). Towards a Theory of OI: Three Core Process Archetypes, *R&D Management, 6*, 1–18.

Gassmann, O., Enkel, E. and Chesbrough, H. (2010). The Future of OI, *R&D Management, 40*(3), 213–221.

Gassmann, O. and Reepmeyer, G. (2005). Organizing Pharmaceutical Innovation: From Science-Based Knowledge Creators to Drug-Oriented Knowledge Brokers, *Creativity and Innovation Management, 14(*3), 233–245.

Gassmann, O., Sandmeier, P. and Wecht, C. H. (2006). Extreme Customer Innovation in the front-end: Learning from a New Software Paradigm, *International Journal of Technology Management, 1*, 46–66.

Gatigon, H. et al. (2002). *A Structural Approach to Assessing Innovation: Construct Development of Innovation Locus, Type and Characteristic. 2002/71/MKT* [online]. Available at: https://flora.insead.edu/fichiersti_wp/inseadwp2002/2002-71.pdf.

Gaule, A. (2006). *OI in Action: How to Be Strategic in the Search for New Sources of Value*. London: Blackwell.

Gault, F. (2018). Defining and measuring innovation in all sectors of the economy. *Research Policy*, *47*(3), 617–622.

Gavetti, G. and Levinthal, D. (2000). Looking Forward and Looking Backward: Cognitive and Experiential Search, *Administrative Science Quarterly, 45*(1), 113–137.

Gawer, A. (2009). *Platforms, Markets and Innovation*. Cheltenham, UK: Edward Elgar Publishing, Ltd.

Gawer, A. and Cusumano, M. A. (2002). *Platform Leadership: How Intel, Microsoft, and Cisco Drive Industry Innovation*. Boston: Harvard Business School Press.

Gawer, A. and Henderson, R. (2007). Platform Owner Entry and Innovation in Complementary Markets: Evidence from Intel, *Journal of Economics & Management Strategy, 16*, 1–34.

Gebauer, H. and Fleisch, E. (2007). An investigation of the Relationship between Behavioral Processes, Motivation, Investments in the Service Business and Service Revenue, *Industrial Marketing Management, 36*(3), 333–348.

Gentile-Lüdecke, S., Torres de Oliveira, R., & Paul, J. (2020). Does Organizational Structure Facilitate Inbound and Outbound OI in SMEs? *Small Business Economics*, *55*, 1091–1112.

Gerlach, M. (1992). *Alliance Capitalism*. Berkeley, CA: University of California Press.

Ghazawneh, A. and Henfridsson, O. (2013). Balacing Platform Control and External Contribution In Third Party Development: The Boundary Resources Model, *Information System Journal, 23*, 173–192.

Ghosh, R. A. (1998). Interview with Linus Torvalds: What Motivates Free Software Developers?, *First Monday - Peer-Reviewed Journal on the Internet, 3*(3).

Gibbons, M. et al. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. London: Sage Publications, Inc.

Gibson, C. A. B. and Birkinshaw, J. (2004). The Antecedents, Consequences, and Mediating Role of Organizational Ambidexterity, *Academy of Management Journal, 47*(2), 209–226.

Gilbert, C. (2005). Unbundling the Structure of Inertia: Resource versus Routine Rigidity, *Academy of Management Journal, 48*(5), 741–763.

GIlfillan, S. C. (1935). *Inventing the Ship*. Chicago: Follett.

Glaser, B. and Strauss, A. (1967). *The Discovery of Grounded Theory*. Hawthorne, NY: Aldine Publishing Company.

Glasser, B. G., & Strauss, A. (2009). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Goes, P. B. (2014). Editor's Comments: Big Data and IS Research. *MIS Quarterly*, *38*(3), iii–viii.

Goldman, S., Nagel, R. and Preiss, K. (1995). *Agile Competitors and Virtual Organizations*. New York, NY: van Nostrand Reinhold.

Gomes de Souza, A. (2014). What is the Role of the Sensitive Construct Theory in Free and Open Source Software Development?, *International Journal of Innovation, Management and Technology, 5*(6), 474–478.

Gomes-Casseres (1996). *The Alliance Revolution: The New Shape of Business Rivalry.* Cambridge, MA: Harvard University Press.

Granovetter, M. S. (1973). The Strenght of Weak Ties, *American Economic Review, 78*, 1360–1380.

Granstrand, O. (2000). *Corporate Innovation Systems: A Comparative Study of Multi-technology Corporations in Japan, Sweden and the USA.* Chalmers University.

Granstrand, O. (2004). The Economics and Management of Technology Marketing: Towards a Pro-Licensing Era?, *International Journal of Technology Management, 27*, 209–240.

Granstrand, O. et al. (1992). External Technology Acquisition in Large Multi-Technology Corporations, *R&D Management, 22*, 111–131.

Greco, M., Grimaldi, M. and Cricelli, L. (2015). OI Actions and Innovation Performance: A Literature Review of European Empirical Evidences, *European Journal of Innovation Management, 2*, 150–171.

Green, S. G., Gavin, M. B. and Aiman-Smith, L. (1995). Assessing a Multidimensional Measure of Radical Technical Innovation, *IEEE Transactions On Engineering Management, 42*(3), 203–213.

Green, J., & Thorogood, N. (2004). *Qualitative methods for health research*. Sage.

Grewal, G. S. et al. (2006). Measurement Science and Technology: A novel approach to thermochromic liquid crystal calibration using neural networks, *Measurement Science and Technology, 17*(7).

Griffith, D. and Pearson, A. W. (1973). The Organisation of Applied Research and Development with Particular Reference to the Customer – Contractor Principle, *R&D Management, 3*, 121–124.

Griffiths, D. . and Pearson, A. W. (1973). The Organisation of Applied Research and Development with Particular Reference to the Customer – Contractor Principle, *R&D Management, 3*, 121–124.

Grinstein, A. (2008). The Relationships Between Market Orientation and Alternative Strategic Orientations: A Meta-analysis, *European Journal of Marketing, 42*(1–2), 115–134.

Gruber, M. and Henkel, J. (2006). New Ventures Based on OI: An Empirical Analysis of Start-Up Firms in Embedded Linux, *International Journal of Technology Management, 33*(4), 356–372.

Guess, A. (2016). *Why Some of the Biggest Names in Big Data are Open Source* [online]. Available at: https://www.dataversity.net/biggest-names-big-data-open-source/

Guest, G., Bunce, A., & Johnson, L. (2006). How Many Interviews Are Enough? An Experiment with Data Saturation and Variability. *Field Methods*, *18*, 59–82.

Guilhon, B., Attia, R. and Rizoulières, R. (2004). Markets for Technology and Firms' Strategies: The Case of the Semiconductor Industry, *International Journal of Technology Management, 27*(2–3).

Gupta, S., Kar, A. K., Baabdullah, A., & Al-Khowaiter, W. A. A. (2018). Big Data with Cognitive Computing: A Review for the Future. *International Journal of Information Management*, *42*, 78–89.

Gupta, A. k., Smith, K. G. and Shalley, C. E. (2006). The Interplay between Exploration and Exploitation, *Academy of Management Journal, 49*(4), 693–706.

Guggenberger, T. M., Möller, F., Haarhaus, T., Gür, I., & Otto, B. (2020). Ecosystem Types in Information Systems. *Guggenberger et al./Ecosystem Types in Information Systems Twenty-Eigth European Conference on IS(ECIS2020)*.

Hacklin, F., Björkdahl, J., & Wallin, M. W. (2018). Strategies for Business Model Innovation: How Firms Reel in Migrating Value. *Long Range Planning*, *51*(1), 82–110.

Hadligton, L. and Scase, M. (2018). End-user Frustrations and Failures in Digital Technology: Exploring the Role of Fear of Missing Out, Internet Addiction and Personality, *Heliyon, 4*(11), 1–18.

Hafkesbrink, J. and Schroll, M. (2014). Ambidextreous Organizational and Individual Competencies in OI: The dawn of a new Research Agenda. *Journal of Innovation Mangement, 2*(1). 9-46.

Hage, J. T. (1980). *Theories of Organizations: Form, Process, and Transformation*. New York, NY: John Wiley & Sons.

Hahn, J., Moon, J. Y. and Zhang, C., J. (2008). Emergence of New Project Teams from Open Source Software Developer Networks: Impact of Prior Collaboration Ties, *ISResearch, 19*.

Hannan, M. T. and Freeman, J. (1989). *Organizational Ecology*. Cambridge, MA: Harvard University Press.

Hannan, M. T. and Freeman, J. (1977). The Population Ecology of Organizations, *American Journal of Sociology, 82*, 929–964.

Hargadon, A. B. and Sutton, R. (1997). Technology Brokering and Innovation in a Product Development Firm, *Administrative Science Quarterly, 42*(4), 716–749.

Harhoff, D., Henkel, J. and von Hippel, E. (2003). Profiting from Voluntary Information Spillovers: How Users Benefit by Freely Revealing Their Innovations, *Research Policy, 23*, 1753–1769.

Hars, A. and Ou, S. (2001). *Working for Free? — Motivations of Participating in Open Source Projects*, in Proceedings of the 34th Hawaii International Conference on System Sciences. Hawaii.

Harte, J. et al. (2001). Business as a Living System: The Value of Industrial Ecology (A Round-table Discussion), *California Management Review, 43*, 16–25.

He, Z.-L. and Wong, P.-K. (2004). Exploration vs. Exploitation: An Empirical Test of the Ambidexterity Hypothesis, *Organization Science, 15*(4), 481–494.

Heidelberger, C. A., & Uecker, T. W. (2009). Scholarly Personal Narrative as ISResearch Methodology. *MWAIS Midwest.*

Helfat, C. E., Finkelstein, S., Mitchell, W., Peteraf, M., Singh, H., Teece, D., & Winter, S. G. (2007). *Dynamic Capabilities: Understanding Strategic Change in Organizations*. Blackwell Publishing .

Helfat, C. E. et al. (2007). *Dynamic Capabilities: Understanding strategic change in organizations*. Oxford: Blackwell.

Helm, R., Endres, H., & Hüsig, S. (2019). When and How often to Externally Commercialize Technologies? A Critical Review of Outbound OI. *Review of Managerial Science*, *13*, 327–345.

Helms, M. M. and Nixon, J. (2010). Exploring SWOT Analysis – Where Are We Now? A Review of Academic Research from the Last Decade, *Journal of Strategy and Management, 3*(3), 215–251.

Henderson, J. C. (1990). Plugging into Strategic Partnerships: The Critical IS Competition, *Sloan Management Review, 31*, 7–18.

Henderson, R. M. (1995). Of Life Cycles Real and Imaginary: The Unexpectedly Long Old Age of Optical Lithography, *Research Policy, 24*, 631–643.

Henderson, R. M. and Clark, K. B. (1990). Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms, *Administrative Science Quarterly, 35*(1), 9–30.

Henderson, R. and Cockburn, I. (1994). Measuring Competence? Exploring Firm Effects in Pharmaceutical Research, *Strategic Management Journal, 15*, 63–84.

Henfridsson, O. and Bygstad, B. (2013). The Generative Mechanisms of Digital Infrastructure Evolution, *MIS Quarterly, 37*, 907–931.

Henkel, J. (2006). Selective Revealing in OI Processes: The Case of Embedded Linux, *Academy of Management Annual Meeting Proceedings, 35*(7), 953–969.

Henkel, J., Schöberl, S. and Alexy, O. (2014). Innovation, The Emergence of Openness: How and Why Firms Adopt Selective Revealing in OI, *Research Policy, 43*, 879–890.

Herriott, R. E. and Firestone, W. A. (1983). Multisite Qualitative Policy Research: Optimizing Description and Generalizability, *Educational Researcher, 12*(2), 14–19.

Hertel, G., Niedner, S. and Herrmann, S. (2003) Motivation of Software Developers in Open Source Projects: An Internet-Based Survey of Contributors to the Linux Kernel, *Research Policy, 32*(7), 1159–1177.

Hertel, G., Niedner, S. and Herrmann, S. (2003). Motivation of Software Developers in Open Source Projects: An Internet-based Survey of Contributors to the Linux Kernel, *Research Policy, 32*, 1159–1177.

Herzog, P. and Leker, J. (2007). *Open vs. Closed Innovation Strategies: Also Different Innovation Cultures?*, in Proceedings of the R&D Management Conference (RADMA). Bremen, Germany.

Hill, S. A. and Birkinshaw, J. (2014). Ambidexterity and Survival in Corporate Venture Units. *Journal of Management*, *40*, 1899–1931.

Hill, C. W. L. and Jones, G. R. (1998). *Strategic Management: An Integrated Approach*. Boston, New York: Houghton Mifflin Company.

Hinchman, L., & Hinchman, S. (1997). Introduction. In L. Hinchman & S. Hinchman (Eds.), *Memory, identity, community: The idea of narrative in the human sciences* (pp. 13–30). State University of New York Press.

Hippel, V. and von Krogh, G. (2003). Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science, *Organization Science, 14*, 209–223.

Hise, R. T. et al. (1990). Marketing/R&D Interaction in New Product Development: Implications for New Product Success Rates, *Journal of Product Innovation Management, 7*(2), 142–145.

Holmqvist, M. (2004). Experiential Learning Processes of Exploitation and Exploration Within and Between organizations: An Empirical Study of Product Development, *Organizational Science, 15*(1), 70–81.

Hou, D. (2007). Studying the Evolution of the Eclipse Java Editor, in *Eclipse: Proceeding OOP-SLA Workshop on Eclipse Technology Exchange*. New York: ACM, (65–69).

Hsiegh, C.-T., Huang, H.-C., & Lee, W.-L. (2016). Using Transaction Cost Economics to Explain OI in Star-ups. *Management Decision* , *54*, 2133–2156.

Hu, Y., McNamara, P. and Mcloughlin, D. (2015). Outbound OI in Bio-Pharmaceutical Out-Licensing, *Technovation, 35*, 46–58.

Hudson, J. (2004). Introduction: Aid and Development, *The Economic Journal, 114*(496), 185–190.

Hudson, P. (2004). Industrial Organization and Structure, in Floud, R. and Johnson, P. (Eds.): *The Cambridge Economic History of Britain. Vol 1, Industrialisation*, 1700-1860. Cambrigde: Cambridge University Press.

Hughes, P. T. (1987). The Evolution of Large Technological System, in Bijker, W. E., Hughes, T. P., and Pinch, T. (Eds.): *The Social Construction of Technological Systems*. Cambrigde, MA.

Hughes, P. T. (1983). *Networks of Power*. Baltimore: Johns Hopkins University Press.

Huizingh, E. K. R. E. (2011). OI: State of the Art and Future Perspectives, *Technovation, 31*(1), 2–9.

Hung, K.-P. and Chou, C. (2013). The Impact of OI on Firm Performance: The Moderating Effects of Internal R&D and Environmental Turbulence, *Technovation, 2013*, 368–380.

Hunt, S. D. (2000). *A General Theory of Competition*. Thousands Oaks, CA.: Sage Publications, Inc.

Hurmelinna, P., Kylaheiko, K. and Jauhiainen, T. (2007). The Janus Face of the Appropriability Regime in the Protection of Innovations: Theoretical Reappraisal and Empirical Analysis, *Technovation, 27*(3), 133–144.

Humphrey, W. S. (1989). *Managing the Software Process*. Boston, MA: Addison-Wesley.

Hustad, T. P. (1991). From the Editor, *The Journal of Product innovation Management, 8*(2), 73–74.

Huston, L. and Sakkab, N. (2006). Connect and Develop Inside Procter & Gamble's New Model for Innovation *Harvard Business Review, 84*(3), 58–67.

Huston, L. and Sakkab, N. (2007). Implementing OI, *Research-Technology Management, 84*(3), 58–67.

Hwang, B., Lai, Y.-P., & Wang, C. (2021). Open Innovation and Organizational Ambidexterity. *European Journal of Innovation Management*, 1–23.

Hwang, J., Kim, E. and Kim, S. (2009). Factors Affecting Open Technological Innovation in Open Source Software Companies in Korea, *Journal Innovation Organization and Management, 11*(3), 279–290.

Iansiti, M., & Levien, R. (2004). *The Keystone Advantage: What the new Dynamics of Business Ecosystems Mean for Strategy, Innovation, and Sustainability*. Harvard Business School Publishing .

Iansiti, M. and Clark, K. (1994). Integration and Dynamic Capability, *Industry and Corporation Change, 3*, 557–606.

Idrissi Fakhredinne, M. O. and Castonguay, Y. (2019). An Exploratory Study on How SMEs are Open to External Sources of Information. *European Journal of Innovation Management*, *22*(5), 765–789.

Ili, S., Albers, A. and Miller, S. (2010). OI in the Automotive Industry, *R&D Management, 40*(3), 246–255.

Inauen, M. and Schenker-Wicki, A. (2012). Fostering Radical Innovations with OI, *European Journal of Innovation Management, 15*(2), 212–231.

Jacob, S. A., & Furgerson, S. P. (2012). Writing Interview Protocols and Conducting Interviews: Tips for Students New to the Field of Qualitative Research. *Qualitative Report*, *17*, 1–10.

Jacobides, M. G. and Billinger, S. (2006). Designing the Boundaries of the Firm: From ”Make, Buy, or Ally” to the Dynamic Benefits of Vertical Architecture, *Organizational Science, 17*(2), 249–261.

Jansen, J. J. P., Van Den Bosch, F. A. J. and Volberda, H. W. (2006). Exploratory Innovation, Exploitative Innovation, and Performance: Effects of Organizational Antecedents and Environmental Moderators, *Management Science, 52*(11), 1661–1674.

Janssen, O. and Huang, X. (2008). Us and Me: Team Identification andIndividual Differentiation as Complementary Drivers of Team Members' Citizenship and Creative Behaviors, *Journal of Management, 34*(1), 69–88.

Jenkins, R. V. (1975). *Images and Enterprise*. Baltimore: Johns Hopkins University Press.

Jeppesen, L. B. and Lakhani, K. R. (2010). Marginality and Problem-Solving Effectiveness in Broadcast, *Organization Science, 21*, 1016–1033.

Jia, F. and Lamming, R. (2013). Cultural Adaptation in Chinese-Western Supply Chain Partnerships: Dyadic Learning in an International Context, *International Journal of Operations & Production Management*, *33*(5), 528–561.

Jones, M. R. (2004). Debatable Advice and Inconsistent Evidence: Methodology in ISResearch, in Kaplan, B. et al. (Eds.) *ISResearch: Relevant Theory and Informed Practice*. Boston: Kluwer Academic Publishers.

Junni, P., Sarala, R. M., Taras, V., & Tarba, S. Y. (2013). Organizational Ambidexterity and Performance: A Meta-Analysis. *Academy of Management Perspectives*, *27*(3), 299–312.

Kahn, K. B. (2018). Understanding Innovation. *Business Horizons*, *61*(3), 453–460.

Kallinikos, A., Aaltonen, A. and Morton, A. (2013). The Ambivalent Ontology of Digital Artifacts, *MIS Quarterly, 37*(2), 357–370.

Kani, M. and Motohashi, K. (2012). Understanding the Technology Markets for Patents: New Insights from a Licensing Survey of Japanese Firms, *Research Policy, 41*, 226-235.

Kanungo, S. (1993). Information Systems: Theoretical Development and Research Approaches. *Information Systems*, *18*(8), 609–619.

Kaplan, S. M. (1999). Discontinuous Innovation and the Growth Paradox, *Strategy & Leadership, 27*(2), 16–21.

Katila, R., Rosenberger, J. D., & Eisenhardt, K. M. (2008). Swimming with Sharks: Technology Ventures, Defense Mechanisms and Corporate Relationships. *Administrative Science Quarterly`*, *53*(2).

Kassotaki, O., Paroutis, S., & Morrell, K. (2019). Ambidexterity Penetration across multiple organizational levels in an aerospace and Defense organization. *Long Range Planning*, *52*(3), 366–385.

Katila, R., & Ahuja, G. (2002). Something old, Something New: Something old, somethA Longitudinal Study of Search Behavior and New Product Introduction. *Academy of Management Journal*, *45*, 1183–1194.

Katz, R. and Allen, T. J. (1982) Investigating the Not Invented Here (NIH) Syndrome: A Look at the Performance, Tenure, and Communication Patterns of 50 R & D Project Groups, *R&D Management, 12*(1), 7–20.

Kauppilla, O.-P. (2010). Creating Ambidexterity by Integrating and Balancing Structurally Separate Interorganizational Partnerships, *Strategic Organization, 8*(4), 283–312.

Kijl, B. and Boersma, D. (2010). *Developing a Business Model Engineering & Experimentation Tool–The Quest for Scalable "Lollapalooza Confluence Patterns"*, in AMCIS 2010 Proceedings. Lima, Peru.

Kim, H.-S. and Choi, S.-Y. (2014). Technological Alliance Portfolio Configuration and Firm Performance. *Review Management Science*, *8*, 541–558.

Kim, H., Kim, Y. and Cho, K. (2014). The Effect of Research and Development Investment and Desorptive Capacity on Firm Performance, *Asian Journal of Technology Innovation, 22*(2), 252–267.

Kim, Y. and Vonortas, N. S. (2006). Determinants of Technology Licensing: The Case of Licensors, *Managerial and Decision Economics, 27*(4), 235-249.

Kim, J., Yoon, J., & Lee, J.-D. (2021). Dominant design and evolution of technological trajectories: The case of tank technology, 1915–1998. *Journal of Evolutionary Economics*, *31*, 661–676.

Kimberly, K. J. (1981). Managerial Innovation, in Nystrom, P. C. and Starbuck, W. H. (Eds.) *Handbook of Organization Design*. Oxford: Oxford University Press.

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

Kleinschmidt, E. J. and Cooper, R. G. (1991). The Impact of Product Innovation on Performance, *Journal of Product Innovation Management, 8*, 240–251.

Klepper, S. (1996). Entry, Exit, Growth, and Innovation over the Product Cycle, *American Economic Review, 86*(3), 562–583.

Klevorick, A. K., Levin, R. C., Nelson, R. R., & Winter, S. G. (1995). On the Sources and Significance of Interindustry Differences in Technological Opportunities. *Research Policy*, *24*(2), 185–205.

Kline, D. (2003). Sharing the Corporate Crown Jewels, *MIT Sloan Management Review, 44*, 89–93.

Kline, S. J. and Rosenberg, N. (1986). An Overview of Innovation, in Landau, R. and Rosenberg, N. (Eds.). *The Positive Sum Strategy.* Washington: National Academy Press (275–304).

Klonglan, G. E. and Coward, E. W. J. (1970). The Concept of Symbolic Adoption: A Suggested Interpretation, *Rural Sociology, 35*(1), 77–83.

Knight, K. E. (1967). A Descriptive Model of Intra- Firm Innovation Process, *Journal of Management, 40*(4), 478–496.

Koch, C. (2004). Innovation Networking Between Stability and Political Dynamics, *Technovation, 24,* 729–739.

Kodoma, F. (1997). *Emerging Patterns of Innovation: Sources of Japan's Technological Edge*. Cambridge, MA: Harvard Business School Press.

Kogut, B. and Zander, U. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology, *Organizational Science, 3*(3), 383–397.

Koruna, S. (2004). External Technology Commercialization: Policy Guidelines, *International Journal of Technology Management, 27*, 241–254.

Koschatzky, K. (2001). Networks in Innovation Research and Innovation Policy — An Introduction, in *Innovation Networks Concepts and Challenges in the European Perspective*. Physica-Verlag.

Kotler, P. and Armstrong, G. (1994). *Principles of Marketing*. Englewood Cliffs, NJ: Prentice Hall.

Kotsemir, M., Abroskin, A. and Meissner, D. (2013). *Innovation Concepts and Typology: An Evolutionary Discussion*. WP BRP 05/STI/2013. Munich.

Krishnamurthy, S. (2006). On the Intrinsic and Extrinsic Motivation of Free/Libre/Open Source (FLOSS) Developers, *Knowledge, Technology and Policy, 18*, 17–39.

Kuhn, T. (1962). *The Structure of Scientific Revolution*. Chicago: University of Chicago Press.

Lakhani, K. R., Lifshitz-Assaf, H. and Tushman, M. (2013). OI and Organizational Boundaries: Task Decomposition, Knowledge Distribution and the Locus of Innovation, in Grandori, A. (Ed.) *Handbook of Economic Organization*. Edward Edgar, (355–382).

Lakhani, K. R. and von Hippel, E. (2003). How Open Source Software Works: Free User to User Assistance, *Research Policy, 32*(6), 923–943.

Lakhani, K. R. and Wolf, R. G. (2005). Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects, in Feller, J. et al. (Eds.). *Free and Open Source Software*. Cambridge: MIT Press.

Lall, S. (1992). Technological Capabilities and Industrialization. *World Development*, *20*(2), 165–186.

Lamoreaux, N. and Sokoloff, K. L. (2001). Market Trade in Patents and the Rise of a Class of Specialized Inventors in the 19th- Century United States, *The American Economic Review, 91*(2), 30–44.

Landau, R. (1984). *The Nature of Technological Knowledge*. Boston: D. Reidel.

Langlois, R. N. and Robertson, P. L. (1992). Networks and Innovation in a Modular System: Lessons from the Microcomputer and Stereo Component Industries, *Research Policy, 21*, 297–313.

Laudan, R. (1984). Cognitive Change in Technology and Science, in *The Nature of Technological Knowledge*. Boston, MA: D. Reidel Publishing Co.

Laursen, K. and Salter, A. (2006). OI: The Role of Openness in Planning Innovation Performance Among U.K. Manufacturing Firms, *Strategic Management Journal, 27*, 131–150.

Laursen, K. and Salter, A. (2004). Searching High and Low: What Types of Firms Use Universities as a Source of Innovation?, *Research Policy, 33*(8), 1201–1215.

Lave, J. (1998). *Communities of Practice*. New York: Cambridge University Press.

Lave, J. and Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Lavie, D. (2006). Capability Reconfiguration: an Analysis of Incumbent Responses to Technological Change. *Academy of Management Review*, *31*, 153–174.

Lavie, D. and Rosenkopf, L. (2006). Balancing Exploration and Exploitation in Alliance Formation, *Academy of Management Journal, 49*(4), 797–818.

Lavie, D., Stettner, U. and Tushman., M. L. (2010). Exploration and Exploitation within and across Organizations, *Academy of Management Annals, 4*(1), 109–155.

Lazzarotti, V. et al. (2017). Innovation Ambidexterity of Open Firms. The Role of Internalrelational Social Capital, *Technology Analysis & Strategic Management, 29*(1), 105–118.

Le Masson, P. et al. (2012). Revisiting Absorptive Capacity with a Design Perspective, *International Journal of Knowledge Management Studies, 5*, 10–44.

Lee, A. (1991). Integrating Positivist and Interpretivist Approaches to Organizational Research, *Organization Science*, 342–365.

Lee, A. S. (1989). A Scientific methodology for MIS case studies, *MIS Quarterly, 13*(1), 33–50.

Lee, D. and Mendelson, H. (2008). Divide and Conquer: Competing with Free Technology under Network Effects, *Production and Operations Management*, 12–28.

Lee, I and Kim, E. (2019). Factors Affecting the Outbound OI Strategies in Pharmaceutical Industry: Focus on Out-Licensing Deal, *Journal of OI: Technology, Market, and Complexity, 5*(4), 1-13.

Lee, K., & Yoo, J. (2019). How Does OI Lead Competitive Advantage? A Dynamic Capability View Perspective. *PloS One*, *14*(11).

Lee, M. and Na, D. (1994). Determinants of Technical Success in Product Development When Innovative Radicalness is Considered, *Journal of Product Innovation Management, 11*, 62–68.

Lee, S.-Y., Kim, H.-W. and Gupta, S. (2009). Measuring Open Source Software Success, *Omega, 37*(2), 426–438.

Lee, S.-Y. et al. (2014). The Green Bullwhip Effect: Transferring Environmental Requirements Along a Supply Chain, *International Journal of Production Economics, 156*(10), 39–51.

Leimeister, J., Huber, M., Bretschneider, U., & Krcmar, H. (2009). Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition. *Journal of Management Information Systems*, *26*(1), 197–224.

Li, R., Fu, L., & Liu, Z. (2020). Does Openness to Innovation Matter? The Moderating Role of OI between Organizational Ambidexterity and Innovation Performance. *Asian Journal of Technology Innovation*, *28*(2), 251–271.

Leonard-Barton, D. (1995). Wellsprings of Knowledge: Building and Sustaining the Sources of Innovation, in Huber, G. P. and Van de Ven, A. (Eds.): *Longitudinal Field Research Methods: Studying Processes of Organizational Change*. Thousand Oaks, CA: Sage.

Lerner, J. and Tirole, J. (2002). Some Simple Economics of Open Source, *Journal of Industrial Economics, 50*(2), 197–234.

Lerner, J. and Tirole, J. (2005). The Economics of Technology Sharing:Open Source and Beyond, *Journal of Economic Perspective, 19*(2), 99–120.

Lesser, E. and Storck, J. (2001). Communities of Practice and Organizational Performance, *IBM Systems Journal, 40*(4), 831–841.

Leten, B. et al. (2013). IP Models to Orchestrate Innovation Ecosystems, *California Management Review, 55,* 51–64.

Lettl., C., Herstatt, C. and Gemuenden, H. G. (2006). Users' Contributions to Radical Innovation: Evidence from four Cases in the Field of Medical Equipment Technology, *R&D Management, 36*(3), 251–272.

Levinthal, D. A. and March, J. G. (1993). The myopia of Learning, *Strategic Management Journal, Winter*(Special Issue), 95–112.

Levitt, B. and March, J. G. (1988). Organizational Learning, *Annual Review of Sociology, 14*, 319–340.

Lewin, A. Y., Long, C. P. and Carroll, T. N. (1999). The Coevolution of New Organizational Forms, *Organization Science, 10*(5), 535–550.

Li, C.-R., Lin, C.-P. and Chu, C.-P. (2008). The nature of Market Orientation and the Ambidexterity of Innovations, *Management Decision, 46*(7), 1002–1026.

Li, J.-B. and Zhang, R. (2020). Summary of Regional Innovation Ecosystem Research. *6th International Conference on Energy Science and Chemical Engineering, IOP Conf. Series: Earth and Environmental Science*.

Li, Y., Tan, C.-H. and Teo, H.-H. (2012). Leadership Characteristics and Developers' Motivation in Open Source Software Development, *Information & Management, 49*(5), 257–267.

Lincoln, Y. S. and Guba, E. G. (1985). *Naturalistic Inquiry*. London: Sage Publications.

Linder, J. and Cantrell, S. (2000). *Changing Business Models: Surveying the Landscape*.

Lindner, J. C., Jarvenpaa, S. and Davenport, T. H. (2003). Toward an Innovation Sourcing Strategy, *MIT Sloan Management Review, 17*, 43–49.

Lindgardt, Z., Reeves, M., Stalk, G., & Deimler, M. S. (2009). *Business Model Innovation. When The Game Gets Tough, Change the Game*. The Boston Consulting Group.

Lorenzi, B. A., Mantel, M. I. and Riley, R. T. (1912). Preparing your Organization for Technological Change, *Healthcare Informatics*.

Lovelock, C. H. and Gummesson, E. (2004). Whither Services Marketing? In Search of a New Paradigm and Fresh Perspectives, *Journal of Service Research, 7*(1), 20–41.

Lowe, A., Norris, A. C., Farris, A. J. and Babbage, D. R. (2018). Quantifying Thematic Saturation in Qualitative Data Analysis. *Field Methods*, *30*(3).

Lubatkin, M. H., Simsek, Z., Ling, Y. and Veiga, J. F. (2016). Ambidexterity and Performance in Small-to Medium-Sized Firms: The Pivotal role of Top Management Team Behavioral Integration, *Journal of Management, 32*(5), 646–672.

Lundvall, B. (1992). *National Systems of Innovation: Towards a Theorem of Innovation and Inter-Active Learning*. London: Pinter.

Lusch, R. F. and Nambisan, S. (2015). Service Innovation: A Service-Dominant Logic Perspective, *MIS Quarterly, 39*, 155–175.

Lusch, R. F., Vargo, S. L. and Wessels (2008). Toward a Conceptual Foundation for Service Science: Contributions from Service-Dominant Logic, *IBM Systems Journal, 47*(1), 5–14.

Lynn, G. S., Morone, J. G., & Paulson, A. S. (1996). Marketing and Discontinuous Innovation: The Probe and Learn Process. *California Management Review*, *38*(3).

Lyytinen, K. and Damsgaard, J. (2010). Configuration analysis of inter-organizational information systems adoption. *Scandinavian Conference on Information Systems*, 127–138.

Lyytinen, K. and Rose, G. M. (2003). The Disruptive Nature of Information Technology Innovations: The Case of Internet Computing in System Development Organizations, *MIS Quarterly, 27*, 557–595.

Lyytinen, K. and Rose, G. M. (2006). Information System Development Agility as Organizational Learning, *European Journal of Information Systems, 15*, 183-199.

MacCormack, A., & Verganti, R. (2003). Managing the Sources of Uncertainty: Matching Process and Context in Software Development. *Journal of Product Innovation Management*, *20*(3), 217–232.

Macdonald, S. (1993). Nothing Either Good or Bad: Industrial Espionage and Technology Transfer. *8*(1–2), 95–105.

Magnusson, T., Lindstrom, G., & Berggren, C. (2003). Architectural or modular innovation? Managing discontinuous product development in response to challenging environmental performance targets. *International Journal of Innovation Management*, *7*(1), 1–26.

Mc Farlan, F. W. (1981). Portofolio Approach to Information Systems, *Harvard Business Review, 59*, 142-150.

Madhavaram, S. and Hunt, S. (2008). The Service-Dominant Logic and a Hierarchy of Operant Resources: Developing Masterful Operant Resources and Implications for Marketing Strategy, *Journal Of The Academy Of Marketing Science, 36*(1), 67–82.

Madjdi, F. and Hüsig, S. (2011). The Heterogeneity of Incumbents' Perceptions and Response Strategies in the Face of Potential Disruptions, *Foresight, 13*, 14–33.

Maglio, P. P. and Spohrer, J. (2008). Fundamentals of Service Science, *Journal of the Academy of Marketing Science, 36*, 18–20.

Maglio, P. P. et al. (2009). The Service System is the Basic Abstraction of Service Science, *ISand e-Business Management, 7*, 395–406.

Majchrzak, A. and Malhotra, A. (2016). Effect of Knowledge-Sharing Trajectories on Innovative Outcomes in Temporary Online Crowds. *ISResearch*, *685*(703), 685–703.

Malerba, F. (2006). Sectoral Systems: How and Why Innovation Differs Across Sectors. In *The Oxford Handbook Of Innovation,*. Oxford University Press.

Malerba, F. and Orsenigo, L. (1996). Schumpeterian Patterns of Innovation are Technology-Specific. *Research Policy*, *25*, 451–478.

Mamasioulas, A., Mourtzis, D., & Chryssolouris, G. (2020). A Manufacturing Innovation Overview: Concepts, Models and Metrics. *International Journal of Computer Integrated Manufacturing*, *33*(8), 769–791.

Mantena, R. and Sankaranarayanan, R. Viswanathan, S. (2007). *Exclusive Licensing in Complementary Network Industries*. NET Institute Working Paper No. 07-04. New York, NY.

March, J. G. (1991). Exploration and Exploitation in Organizational Learning, *Organization Science, 2*(1), 71–87.

Margaret, L. S. and Nathaniel, N. H. (2019). Impact of subsidiaries' cross-border knowledge tacitness shared and social capital on MNCs' explorative and exploitative innovation capability. *Journal of International Management*, *25*(4), 1–16.

Markowitz, H. (1952). Portofolio Selection, *The Journal of Finance,* 7, 77-91.

Markus, M. L. (1983). Power, Politics and MIS Implementation, *Communications of the ACM, 26(*6), 1983.

Marquis, C. and Tilcsik, A. (2013). Imprinting: Toward a Multilevel Theory. *The Academy of Management Annals*, *7*(1), 193–243.

Marshall, A. (1932). Elements of Economics, in *Elements of Economics of Industry*. 3rd ed. London: Macmillan.

Marshall, B., Cardon, P., Poddar, A. and Fontenot, R. (2013). Does sample size matter in qualitative research?: a review of qualitative interviews in IS research. *Journal of Computer Information Systems*, *54*(1), 11–22.

Marshall, M. N. (1996). Sampling for Qualitative Research, *Family Practice, 13*(6), 522–525.

Martin, B. R. (2016). Twenty Challenges for Innovation Studies. *Science and Public Policy*, *43*(3), 432–450.

Maskell, P. et al. (1998). *Competitiveness, Localized Learning and Regional Development: Specialization and Prosperity in Small Open Economies*. London: Routledge.

Masucci, M., Brusoni, S. and Cennamo, C. (2020). Removing Bottlenecks in Business Ecosystems: The Strategic Role of Outbound OI, *Research Policy, 49*(1).

Matzler, K., Bailom, F., Friedrich von den Eichen, S., & Kohler, T. (2013). Business Model Innovation: Coffee Triumphs for Nespresso. *Journal of Business Strategy*, *34*(2), 30–37.

Mazzola, E., Bruccoleri, M. and Perrone, G. (2012). The Effect of Inbound, Outbound and Coupled Innovation on Performance, *International Journal of innovation Management, 16*(6), 1-27.

McCracken, G. (1988). *The Long Interview*. Sage Publications.

McGrath, R. G. (2001). Exploratory Learning, Innovative Capacity, and Managerial Oversight, *Academy of Management Journal, 44*(1), 118–131.

McGrath, R. G., Macmillan, I. C. and Tushman, M. (1992). The Role of Executive Team Actions in Shaping Dominant Designs: Towards the Strategic Shaping of Technological Progress, *Strategic Management Journal, 13*(Issue S2: Special Issue), 137–161.

McKelvey, B. (2004). Toward a Complexity Science of Entrepreneurship, *Journal of Business Venturing, 19*, 313–341.

McKusick, K. (1999). Twenty Years of Berkeley Unix, in DiBona, C., Ockman, S., and Stone, M. (Eds.): *Open Sources: Voices from the Open Source Revolution*. Sebastopol, CA: O'Reilly.

Meinlschmidt, J., Foerstl, K. and Kirchoff, H. F. (2016). The Role of Absorptive and Desorptive Capacity in Sustainable Supply Management, *International Journal of Physical Distribution & Logistics Management, 26*(2), 117–211.

Meissner, D., & Shmatko, N. (2017). Keep Open: The Potential of Gatekeepers for the Aligning Universities to the New Knowledge Triangle. *Technological Forecasting and Social Change*, *123*, 191–198.

Meissner, D., & Shmatko, N. (2018). Integrating Professional and Academic Knowledge: The Link Between Researchers Skills and Innovation Culture. *Journal of Technology Transfer*, *44*, 1273–1289.

Mens, T., Fernández-Ramil, J. and Degrandsart, S. (2008). The Evolution of Eclipse, *Proceeding 24th International Conference on Software Maintenance (ICSM)*, 386–395.

Mensch, G. (1979). *Stalemate in Technology*. Cambridge: Ballinger.

Mention, A.-L. (2011). Co-operation and Co-opetition as OI Practices in the Service Sector: Which Influence on Innovation Novelty? *Technovation*, *31*(1), 44–53.

Mergel, I. (2015). Opening Designing OI Processes to Collaborate with External Problem Solvers. *Computer Review* , *33*(5), 599–612.

Merlo, E. et al. (2004). Linear Complexity Object-oriented Similarity for Clone Detection and Software Evolution Analyses, in *20th IEEE International Conference on Software Maintenance Proceedings*. Chicago, IL.

Merton, R. (1968). *Social Theory and Social Structure*. New York: Free Press.

Merriam-Webster. (2022). *Innovation*. https://www.merriam-webster.com/dictionary/innovation

Metcalfe, C. J. (2011). Persistence of Technological Leadership: Emerging Technologies and Incremental Innovation, *The Journal of Industrial Economics, 59*(2), 199–223.

Metcalfe, M. (2005). Generalisation: Learning Across Epistemologies. *Forum: Qualitative Social Research*, *6*(1).

Metcalfe, J. S. (1995). Technology Systems and Technology Policy in an Evolutionary Framework, *Cambridge Journal of Economics, 19*(1), 25–46.

Metcalfe, J. S. (1994). The Economics of Evolution and the Economics of Technology Policy, *Economic Journal, 104*, 931–944.

Meyer, A., Brooks, G. and Goes, J. (1990). Environmental Jolts and Industry Revolutions: Organizational Responses to Discontinuous Change, *Strategic Management Journal, 11*, 93–110.

Meyers, P. W., Tucker, W. and Gaither, F. (1989). Defining Roles for Logistics During Routine and Radical Technological Innovation, *Journal Of The Academy Of Marketing Science, 17,* 73–82.

Michelino, F. et al. (2014). Inbound and Outbound OI: Organization and Performances, *Journal of Technology Management and innovation, 9*(3), 65–82.

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

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Sage Publications.

Miles, R. E., Snow, C. C., Meyer, A. D., & Coleman Jr, H. J. (1978). Organizational Strategy, Structure, and Process. *The Academy of Management Review*, *3*(3), 546–562.

Milgram, S. (1967). The Small World Problem, *Psychology Today, 1*(May), 61–67.

Miller, D. (1986). Configurations of Strategy and Structure: Towards a Synthesis. *Strategic Management Journal*, *7*(3), 233–249.

Mingers, J. (2001). Combining IS Research Methods: Towards a Pluralist Methodology, *ISResearch, 12*(3), 240–259.

Mishler, E. (1995). Models of narrative analysis: A typology. *Journal of Narrative and Life History*, *5*(2), 87–123.

Mitchell, D. W., & Coles, C. (2003). The Ultimate Competitive Advantage of Continuing Business Model Innovation. *Journal of Business Strategy*, *24*(5), 15–21.

Mitev, N. N., & Bartis, E. (2008). A Multiple Narrative Approach to ISFailure: A Successful System that Failed. *European Journal of Information Systems*, *April*, 1421–1432.

Modi, S. and Mabert, V. (2007). Supplier Development: Improving Supplier Performance Through Knowledge Transfer, *Journal of Operations Management, 25*(1), 42–64.

Moen, T. (2006). Reflections on the Narrative Research Approach. *International Journal of Qualitative Methods* , *5*(4), 56–68.

Mohamad Hashim, H. N., Muhamad Khair, M. H., Mahmood, A., & Mohd Zakuan, Z. Z. (2020). Outbound OI Policy for Exploitation of Intellectual Creation, Design and Creativity in Malaysian Public Universities. *Asian Journal of University Educatio*, *16*(4), 44–52.

Moon, J. Y. and Sproull, L. (2002). Essence of Distributed Work: The Case of the Linux Kernel', in Hinds, P. and Kiesler, S. (Eds.): *Distributed Work*. Cambridge: MIT Press.

Moore, F. M. (1993). Predators and Prey: A new Ecology of Competition, *Harvard Business Review, May-June*, 75–86.

Morakanyana, R., Grace, A. and O'Reilly, P. (2017). Conceptualizing Digital Transformation in Business Organizations: A Systematic Review of Literature, in *30th E-Conference: Digital Transformation - From Connecting Things to Transforming out Lives*. Bled, Slovenia.

Morison, E. E. (1966). *Men, Machines, and Modern Times*. Cambrigde, MA: MIT Press.

Mortara, L. and Minshall, T. (2011). How Do Large Multinational Companies Implement OI?, *Technovation, 31*, 586–597.

Motzek, R. (2007). *Motivation in OI: An Exploratory Study on User Innovators*. Saarbrücken: VDM Verlag Dr. Müller.

Mowery, D. C. (1983). The relationship between intrafirm and contractual forms of industrial research in American manufacturing, *Explorations in Economic History, 67*, 297–306.

Mowery, D. C. (2009). Plus Ça change: Industrial R&D in the "Third Industrial Revolution", *Industrial and Corporate Change, 18*(1), 1–50.

Mowery, D. C. (1983). Industrial Research and Firm Size, Survival, and Growth in American Manufacturing, 1921-1946: An Assessment, *The Journal of Economic History, 43(*4), 953–980.

Mumford, E. et al. (1985). *Research Methods in Information Systems*, in IFIP 8.2 Proceedings. North-Holland, Amsterdam.

Munir, H. et al. (2018). OI using open source tools: a case study at Sony Mobile, *Empir Software Eng, 23*, 186–223.

Munson, F. C. and Pelz, D. C. (1979). *The Innovating Process: A Conceptual Framework*. Detroit.

Murmann, J. P., & Frenken, K. (2006). Toward a systematic framework for research on dominant designs, technological innovations, and industrial change. *Research Policy*, *35*, 925–952.

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

Myers, S. and Marquis, D. C. (1969). *Successful Industrial Innovations*. Washington, DC: National Science Foundation, NSF 69- 17.

Mytelka, L. K. and Smith, K. (2002). Policy Learning and Innovation Theory: An Interactiveand Co-Evolving Process, *Research Policy, 31*, 1467–1479.

Nagaoka, S. and Kwon, H. U. (2006). The Incidence of Cross-Licensing: A Theory and new Evidence on the Firm and Contract Level Determinants, *Research Policy, 35*, 1347–1361.

Nakakoji, K., Ye, Y. and Yamamoto, Y. (2010). Comparison of Coordination Communication and Expertise Communication in Software Development: Motives, Characteristics, and Needs, in K., N., Y., M., and E., M. (Eds.): *New Frontiers in Artificial Intelligence: Lecture Notes in Computer Science*, vol 6284. Berlin: Springer.

Napier, N. P., Mathiassen, L. and Johnson, R. D. (2006). *Negotiating Response-Ability and Repeat-Ability in Requirements Engineering*, in Proceedings of International Conference on Information Systems. Milwaukee, WI.

Napier, N. P., Mathiassen, L. and Robey, D. (2011). Building Contextual Ambidexterity in a Software Company to Improve Firm-Level Coordination, *European Journal of Information Systems, 20*(6), 674-690.

Naqshbandi, M. M., Tabche, I., Choudhary, N. (2018). Managing OI: The Roles of Empowering Leadership and Employee Involvement Climate, *Management Decision, 57*(3), 703-723.

Naqshbandi, M., Singh, S. K. G., & Ma, P. (2016). The Link Between Organisational Citizenship Behaviours and OI: A Case of Malaysian High-Tech Sector. *IJMB Management Review*, *28*(4), 1–12.

Nambisan, S. (2013). Information Technology and Product/Service Innovation: A Brief Assessment and Some Suggestions for Future research, *Journal of the Association of Information Systems, 14,* 215–226.

Narula, R. (2001). Choosing Between Internal and Non-Internal R&D Activities: Some Technological and Economic Factors, *Technology Analysis & Strategic Management, 13*(3), 365–387.

Natalicchio, A. et al. (2017). Managing Knowledge Assets for OI: A Systematic Literature Review, *Journal of Knowledge Management, 21*(6), 1362–1383.

Neely, A. (2008). Exploring the Financial Consequences of the Servitization of Manufacturing, *Operations Management Research, 1*, 103–118.

Nelson, R. R. (1993). *National Innovation Systems: A Comparative Analysis*. New York: Oxford University Press.

Nelson, R. R. and Winter, S. (1982). *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press.

Nonaka, I. and Takeuchi, H. (1995). *The Knowledge Creating Company*. Oxford: Oxford University Press.

Nooteboom, B. (1999*). Inter-Firm Alliances: Analysis and Design*. London: Routledge.

Normann, R. (2001). *Reframing Business: When the Map Changes the Landscape*. Chichester: Wiley.

Normann, R. (2002). *Service management; strategy and leadership in service business* (3rd ed.). Wiley.

Nyström, H. (1985). Product Development Strategy, *Journal of Product Innovation Management, 2,* 25–33.

O'Connor, G. C. (1998). Market Learning and Radical Innovation: A Cross Case Comparison of Eight Radical Innovation Projects, *The Journal of Product innovation Management Innovation, 15*(2), 151-166.

O'Mahony, S. (2003). Guarding the Commons: How Community Managed Software ProjectsProtect Their Work, *Research Policy, 32*(7), 1179–1198.

O'Mahony, S. and Ferraro, F. (2007). The Emergence of Governance in an Open Source Community, *Academy of Management Journal, 50*(5), 1079–1106.

O'Reilly, C. A. and Tushman, M. (2004). The Ambidextrous Organization, *Harvard Business Review, 82*(4), 74–81.

O'Reilly, C. A. and Tushman, M. L. (2008). Ambidexterity as a Dynamic Capability: Resolving the Innovators' Dilemma, *Research in Organizational Behavior, 28*, 185–206.

O'Reilly, C. A. and Tushman, M. L. (2013). Organizational Ambidexterity: Past, Present, and Future, *Academy of Management Perspectives, 27*(4), 324–338.

O'Reilly, C. A. and Tushman, M. L. (2007). *Ambidexterity as a Dynamic Capability: Resolving the Innovator's Dilemma.* Working Paper # 07-088. Cambridge, MA.

O'Reilly, M., & Parker, N. (2012). Unsatisfactory saturation": A critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research*, *13*, 190–197.

O'Sullivan, D. and Dooley, L. (2008). *Applying Innovation*. London: Sage Publications.

Oerlemans, L. A. G., Knoben, J., Pretorius, M. W. and Marthinus W. (2013). Alliance portfolio diversity, radical and incremental innovation: the moderating role of technology management. *Technovation*, *33*(6/7), 234–246.

Oke, A., Burke, G. and Myers, A. (2007). Innovation Types and Performance in Growing UK SMEs, *International Journal of Operations & Production Management, 27*, 735–753.

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

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

Orlikowski, W. J. and Iacono, C. S. (2001). Desperately seeking the "IT" in IT research: A call to theorizing the IT artifact, *ISResearch, 12*, 121–134.

Ortiz de Guinea, A. and Raymond, L. (2019). Improving SMEs' Service Innovation Performance in the Face of Uncertainty Through IT Ambidexterity: A Configurational Approach, *Proceedings of the 52nd Hawaii International Conference on System Sciences. Hawaii*, 5518–5527.

Osterwalder, A. and Pigneur, Y. (2009). *Business Model Generation. A Handbook for Visionaries, Game Changers, and Challengers*. Amsterdam: Moddermann Druckwerk.

Osterwalder, A., Pigneur, Y. and Tucci, C. (2005). Clarifying Business Models: Origins Present, and Future of the Concept, *Communications of the Association for Information Systems, 16,* 1–25.

Pagani, M. (2013). Digital Business Strategy and Value Creation: Framing the Dynamic Cycle of Control Points. *MIS Quartely*, *37*(2), 617–632.

Pagell, M., Wu, Z. and Wasserman, M. E. (2010). Thinking Differently about Purchasing Portfolios: An Assessment of Sustainable Sourcing, *Journal of Supply Chain Management, 46(*1), 57–73.

Pan, S. L. and Leidner, D. E. (2003). Bridging Communities of Practice with Information Technology in Pursuit of Global Knowledge Sharing, *Journal of strategic Information Systems, 12*, 71–88.

Papadopoulos, T. et al. (2013). From Open Source to OI Practices: A Case in the Greek Context in Light of the Debt Crisis, *Technological Forecasting and Social Change, 80*, 1232–1246.

Parida, V., Westerberg, M. and Frishammar, J. (2012). Inbound OI Activities in High-Tech Smes: The Impact on Innovation Performance, *Journal of Small Business Management, 50*(2), 283–309.

Park, W.-Y., & Tangpong, C. (2021). Performance and survival implications of sourcing choice sequence across an architectural innovation life cycle. *Journal of Operations Management*, *67*(6), 656–679.

Parker, G. and van Alstyne, M. W. (2008). Managing Platform Ecosystems, *Proceedings of the 29th International Conference on Information Systems*. Paris, France.

Paul, S. et al. (2010). How to Improve R&D Productivity: The Pharmaceutical Industry's Grand Challenge, *Nat. Rev. Drug Discovery, 9*, 203–214.

Pavitt, K. (1984). Sectoral Patterns of Technical Change: Towards Taxonomy and a Theory, *Research Policy, 13*, 343–373.

Pearson, A. W., Green, T. and Ball, D. F. (1979). A Model for Studying some Organizational Effects of an Increase in the Size of R&D Projects, *IEEE Transactions On Engineering Management, 26*(1), 14–21.

Peltz, C. (2003). Web Services Orchestration and Choreography, *Computer, 36*, 46–52.

Peng, D. X., Verghese, A., Shah, R., & Schroeder, R. G. (2013). The relationships between external integration and plant improvement and innovation capabilities: The moderation effect of product Clockspeed. *Journal of Supply Chain Management*, *3*(3–24).

Perens, B. (1999). The Open Source Definition, in Di Bona, C., Ockman, S., and Stone, M. (Eds.): *Open Sources: Voices from the Open Source Revolution*. Boston, MA: O'Reilly, (171–189).

Perez, C. (2012). *Innovation Systems and Policy: Not only for the Rich?* Tallinn.

Perroux, F. (1950). Economic Space: Theory and Applications, *The Quarterly Journal of Economics, 64(*1), 89–104.

Petrova, E. (2014). *Innovation and Marketing in the Pharmaceutical Industry*. New York, NY: Springer.

Pfeffer, J. (1981). *Power in Organizations*. Cambridge: Ballinger.

Pfeffer, J. and Salancik, G. R. (1978) *The External Control of Organizations: A Resource Dependence Perspective*. Stanford, CA: Stanford Business Classics.

Phillips, W., Noke, H., Bessant, J., & Lamming, R. (2006). Beyond the steady state: Managing discontinuous product and process innovation. *International Journal Innovation and Management*, *10*, 175–196.

Picaud-Bello, K., Johnsen, T., Calvi, R., & Giannakis, M. (2019). Exploring early purchasing involvement in discontinuous innovation: A dynamic capability perspective. *European Journal of Purchasing and Supply Management*, *25*(4).

Pierenkemper, T. and Tilly, R. (2004). *The German Economy During the Nineteenth Century*. B. Books: New York.

PPiezunka, H., & Dahlander, L. (2015). Distant search, narrow attention: How Crowding Alters Organizations' Filtering of Suggestions in Crowdsourcing. *Academy of Management Journal*, *58*(3), 856-880.

Piller, F. T. and Walcher, D. (2006). Toolkits for Idea Competitions: A Novel Method to Integrate Users in New Product Development, *R&D Management, 36*(3), 307–318.

Pinch, T. and Bijker, W. E. (1987). The Social Construction of Facts and Artifact, in Bijker, W. E., Highes, T. P., and Pinch, T. (Eds.): *The Social Construction of Technological System. Cambridge*, MA: MIT Press.

Piore, M. J. and Sabel, C. F. (1984). *The Second Industrial Divide*. New York: Basic Books.

Pisano, G. P. (2015). You Need an Innovation Strategy*, Harvard Business Review, June* [online]. Available at: https://hbr.org/2015/06/you-need-an-innovation-strategy

Pisano, G. P. and Teece, D. J. (2007). How to Capture Value from Innovation: Shaping Intellectual Property and Industrial Architecture, *California Management Review, 50*(1), 277-296.

Plerenkmper, T. and Tilly, R. (2004). *The German Economy During the Nineteenth Century*. New York: Berghahn Books.

Plessis, M. D. (2007). The Role of Knowledge Management in Innovation, *Journal of Knowledge Management, 11*(4), 20–29.

Polit, D. F., & Beck, C. T. (2010). *Essentials of Nursing Research: Appraising Evidence for Nursing Practice* (7th ed.). Wolters Kluwer Health/Lippincott Williams & Wilkins.

Popadiuk, S., Souza Luz, A. R., & Kretschmer, C. (2018). Dynamic Capabilities and Ambidexterity: How are These Concepts Related? *Revista de Administração Contemporânea* , *22*(5), 640–660.

Pope, C. and Mays, N. (2006). Qualitative Methods in Health Research, in Pope, C. and Mays, N. (Eds.): *Qualitative Research in Health Care*. Blackwell Publishing (1–11) [online]. Available at: https://psycnet.apa.org/record/2006-12201-001.

Poppers, K. R. (1963). *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge & K. Paul.

Porter, M. E. (1980). *Competitive Strategy*. New York: The Free Press.

Porter, M. E. (1985). *Competitive Advantage*. New York: The Free Press.

Porter, M. E. (1990).*The Competitive Advantage of Nations*. New York: Free Press.

Porter, M. E. (2001). Strategy and the Internet, *Harvard Business Review, 79*, 62–79.

Powell, W. W. and Brantley, P. (1992). Competitive Cooperation in Biotechnology: Learning through Network, in Nohira, N. and Eccles, R. G. (Eds.): *Networks Organizations*. Boston, MA: Harvard Business Review Press.

Powell, W. W., Koput, K. W. and Smith-Doerr, L. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology, *Administrative Science Quarterly, 41*, 116–145.

Prahalad, C. K. and Hamel, G. (1994). Strategy as a Field of Study: Why Search for a new Paradigm?, *Strategic Management Journal, 15*, 5–16.

Prahalad, C. K. and Ramaswamy, V. (2004). Co-Creation Experiences: The Next Practice in Value Creation, *Journal of Interactive Marketing, 18*(3), 5–14.

Prügl, R. and Schreier, M. (2006). Learning from Leading-Edge Customers at The Sims: Opening up the Innovation Process Using Toolkits, *R&D Management, 36*, 237–250.

Puranam, P. and Srikanth, K. (2007). What They Know vs. What They Do: How Acquirers Leverage Technology Acquisitions, *Strategic Management Journal, 28*(8), 805–826.

Quinn, R. E. and Cameron, K. S. (1988). *Paradox and Transformation: Toward a Theory of Change in Organization and Management*. Cambrigde, MA: Ballinger Publishing Co.

Rachinger, M. (2018). Technological Disruption and Business Model Innovation in the Context of the Business Ecosystem: An Investigation of Technologies for Electric

Mobility in the Automotive Industry. *Event 24. Techno-Ökonomie Kolloquium - Technische Universität Wien | Institut Für Managementwissenschaften*.

Rachinger, M., Rauter, R., Müller, C., Vorraber, W., & Schirgi, E. (2019). Digitalization and its Influence on Business Model Innovation. *Journal of Manufacturing Technology Management*, *30*(8), 1143–1160.

Rai, A. and Tang, X. (2014). IT-Enabled Business Models: A Conceptual Framework and a Coevolution Perspective for Future Research, *ISResearch, 25*, 1–14.

Raisch, S., & Birkinshaw, J., Probst, G. and Tushman, M. (2009). Organizational Ambidexterity: Balancing Exploitation and Exploration for Sustained Performance, *Organization Science, 20*(4), 685–695.

Randhawa, K., Wilden, R., & Hohberger, J. (2016). A Bibliometric Review of OI: Setting a Research Agenda. *Journal of Product Innovation Management*, *33*(6), 750–772.

Raisch, S., & Birkinshaw, J. (2008). Organizational Ambidexterity: Antecedents, Outcomes, and Moderators. *Journal of Management*, *34*(3), 375–409.

Raymond, E. S. (1999). *The Cathedral and the Bazaar: Musing on Linux and Open Source by an Accidental Revolutionary*. Cambridge: O'Reilly.

Reed, R., Storrud-Barnes, S., & Jessup, L. (2012). How OI Affects the Drivers of Competitive Advantage: Trading the Benefits of IP Creation and Ownership for Free Invention. *Management Decision*, *50*(1), 58–73.

Remneland Wikhamn, B. and Styre, A. (2019). Managerial Challenges of Outbound OI: A Study of Spinout Initiative in AstraZeneca, *R&D Management, 49*(4), 652-667.

Rescher, N. (1979). *Cognitive Systematization: A Systems-Theoretic Approach to a Coherentist Theory of Knowledge*. Basil Blackwell.

Rice, M. P., Leifer, R. and O'Connor, G. C. (2002). Commercializing Discontinuous Innovations: Bridging the Gap From Discontinuous Innovation Project to Operations, *IEEE Transactions On Engineering Management, 49(*4), 330–340.

Richardson, J. (2008). The Business Model: An Integrative Framework for Strategy Execution, *Briefings in Entrepreneurial Finance, 17*(5–6), 133–144.

Riessman, C. K. (1993). *Narrative Analysis*. Sage Publications.

Ritala, P., Agouridas, V., Assimakopoulos, D., & Gies, O. (2013). Value Creation and Capture Mechanisms in Innovation Ecosystems. *International Journal of Technology Management*, *63*(3–4), 244–267.

Rivette, K. G. and Kline, D. (2000). *Rembrandts in the Attic: Unlocking the Hidden Value of Patents*. Boston, MA: Harvard Business School Press.

Rivkin, J. and Siggelkow, N. (2003). Balancing Search and Stability: Interdependencies Among Elements of Organizational Design, *Manage Science, 49*(3), 290–311.

Roberts, J. A., Hann, I.-H. and Slaughter, S. A. (2006). Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects, *Management Science, 52*(7).

Robertson, T. S. (1967). The Process of Innovation and the Diffusion of Innovation, *Journal of Marketing, 31*(1), 14–19.

Robertson, P. L., Casali, G. L., & Jacobson, D. (2012). Managing open incremental process innovation: absorptive capacity and distributed learning. *Research Policy*, *41*(5), 822–832.

Rogers, E. M. (1995). *Diffusion of Innovations*. New York: The Free Press.

Röglinger, M. Schwindenhammer, L. and Stelzl, K. (2018). *How to Put Organizational Ambidexterity into Practice-Towards a Maturity Model*, in International Conference on Business Process Management. Sydney, Australia, (194–210).

Roldán Bravo, M. I., Ruiz-Moreno, A., & Lloréns Montes, F. J. (2018). Examining desorptive Capacity in Supply Chains: The Role of Organizational Ambidexterity. *International Journal of Operations & Production Management*, *38*(2), 534–553.

Rondé, P., & Hussler, C. (2005). Innovation in Regions: What Does Really Matter? *Research Policy*, *34*(8), 1150–1172.

Romanelli, E. and Tushman, M. (1994). Organizational Transformation as Punctuated Equilibrium: An Empirical Test, *The Academy of Management Journal, 37*(5).

Rosen, M. (1991). Coming to Terms with the Firms: Understanding and Doing Organizational Ethnography, *Journal of Management Studies, 28*(1), 1–24.

Rosenberg, N. (1982). *Inside the Black Box: Technology and Economics*. Cambridge: Cambridge University Press.

Rosenkopf, L. and Nerkar, A. (1999). On the Complexity of Technological Evolution: Exploring Coevolution within and across Hierarchical Levels in Optical Disk Technology, in

Baum, J. and McKelvey, W. (Ed.): *Variations in Organization Science: In Honor of D. T. Campbell*. Thousand Oaks, CA: Sage Publications.

Rosenkopf, L. and Nerkar, A. (2001). Beyond Local Search: Boundary-Spanning Exploration, and Impact in the Optical Disk Industry, *Strategic Management Journal, 22*, 287–306.

Rosenkopf, L. and Tushman, M. (1998). The Coevolution of Community Networks and Technology: Lessons From the Flight Simulation Industry, *Industrial and Corporate Change, 7*, 311–346.

Ross, J. W., Beath, C. M. and Goodhue, D. L. (1996). Develop Long-Term Competitiveness through IT Assets, *Sloan Management Review, 38*, 31–45.

Rothaermel, F. T. and Deeds, D. L. (2004). Exploration and Exploitation Alliances in Biotechnology: A System of New Product Development, *Strategic Management Journal, 25*(3), 201–221.

Rothwell, R. (1974). Factors for Success in Industrial Innovation, *Journal of General Management, 2*(2), 57–65.

Rothwell, R. and Zegveld, W. (1985). *Reindustrialisation and Technology*. London: Longman.

Rouse, W. B. (1994). Strategies for Innovation, *The Engineering Economit, 39*(2).

Rowley, J., Baregheh, A. and Sambrook, S. (2011). Towards an Innovation-Type Mapping Tool, *Management Decision, 49*(1), 73–86.

Russo, N. L. and Stolterman, E. (2000). Exploring the Assumptions Underlying ISMethodologies. Their Impact on Past, Present and Future ISM Research, *Information Technology & People, 13*(4), 313–327.

Ryu, H. S. and Lee, J. N. (2018). Understanding the Role of Technology in Service Innovation: Comparison of Three Theoretical Perspectives, *Information & Management, 55*, 294–307.

Sagiroglu, S., & Sinanc, D. (2003, May 20). Big Data: A Review. *International Conference on Collaboration Technologies and Systems (CTS)*.

Sahal, D. (1981). *Patterns of Technological Innovation*. Reading, MA: Addison-Wesley.

Sahal, D. (1985). Technological Guideposts and Innovation Avenues, *Research Policy, 14*, 61–82.

Salampasis, D. and Mention, A-L. (2019). From A-Value to Value-Multiplication: Leveraging Outbound OI Practices for Unrelated Diversification in the Sensor Industry, *Technology Analysis and Strategic Management, 31*(11), 1327-1340.

Sambamurthy, V., Bharadwaj, A. and Grover, V. (2003). Shaping Agility through Digital Options: Reconceptualizing the Role of Information Technology in Contemporary Firms, *MIS Quarterly, 27*(2), 237–263.

Sanderson, S. and Uzumeri, M. (1995). Managing Product Families: The Case of the Sony Walkman, *Research Police*, 761–782.

Santos, F. M. and Eisenhardt, K. M. (2005). Organizational Boundaries and Theories of Organization, *Organization Science, 16*(5), 491–508.

Santos, J., Spector, B., & Van der Heyden, L. (2009). Toward a Theory of Business Model Innovation within Incumbent Firms. *INSEAD Working Paper Collection*, *INSEAD.*

Sastry, M. A. (1997). Problems and Paradoxes in a Model of Punctuated Organizational Change, *Administrative Science Quarterly, 42*, 237–75.

Saviotti, P. (1996). *Technology Evolution, Variety and the Economy*. Cheltenham, UK: Edward Elgar Publishing.

Saviotti, P. and Metcalfe, J. S. (1991). *Evolutionary Theories of Economic and Technological Change: Present Status and Future Prospects*. Reading: Harwood Academic.

Saviotti, P. and Metcalfe, J. S. (1984). A Theoretical Approach to Construction of Technological Output Indicators, *Research Policy, 13*, 141–151.

Saxenian, A. (1991). The Origins and Dynamics of Production Networks in Silicon Valley, *Research Policy, 20*, 423–437.

Say, J.-B. (1821). *A Treatise on Political Economy. Reprints of Economic Classics*. New York, NY: Augustus M. Kelly.

Scacchi, W. (2004). Free and Open Source Development Practices in the Game Community, *IEEE Software, 21*(1), 59–66.

Schall, E. J. D., Ospina, S., Godsoe, B., & Dodge, J. (2004). Appreciative Narratives as Leadership Research: Matching Method to Lens. *Constructive Discourse and Human Organization Advances in Appreciative Inquiry*, *1*, 147–170.

Schallmo, D., & Brecht, L. (2010). Business Model Innovation in Business-to-Business Markets − Procedure and Examples”. *Proceedings of the 3rd ISPIM Innovation Symposium, Managing the Art of Innovation: Turning Concepts into Reality*, 1–24.

Scherrer-Rathje, M., Deflorin, P. and Anand, G. (2014) ‘Manufacturing Flexibility through Outsourcing: Effects of Contingencies, *International Journal of Operations & Production Management, 9*, 1210–1242.

Schilling, M. A. (2000). Toward a General Modular Systems Theory and its Application to Interfirm Product Modularity, *The Academy of Management Review, 25*(2), 312–334.

Schmookler, J. (1966). *Invention and Economic Growth*. Cambridge, MA: Harvard University Press.

Schroll, A. and Mild, A. (2011). OI Modes and the Role of Internal R&D: An Empirical Study on OI Adoption in Europe, *European Journal of Innovation Management, 14*(4), 475–495.

Schulze, A., Brojerdi, G. and Von Krogh, G. (2014). Those Who Understand, Teach: Disseminative Capability and Knowledge Transfer in the Automotive Industry, *Journal of Product Innovation Management, 31*(1), 79–97.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*. Oxon: Routledge.

Schumpeter, J. A. (1934). *The Theory of Economic Development: An Inquiry Into Profits, Capital, Credit, Interest, and the Business Cycle*. Transaction: New Brunswick.

Schwandt, T. . A. (1994). Constructivist, Interpretivist Approaches to Human Inquiry, in Denzin, N. K. and Lincoln, Y. S. (Eds.): *Handbook of Qualitative Research*. Thousand Oak, CA: Sage Publications.

Selander, L., Henfridsson, O. and Svahn, F. (2010). Transforming Ecosystem Relationship in Digital Innovation, in *International Conference on Information System (ICIS)*. St. Louis, MO, USA.

Sen, T. K. and Ghandforoush, P. (2011). Radical and Incremental Innovation Preferences in Information Technology: An Empirical Study in an Emerging Economy, *Journal of Technology Management and innovation, 6*(4), 33–44.

Sen, R., Subramaniam, C. and Nelson, M. L. (2011). Open Source Software Licenses: Strong-Copyleft, Non-Copyleft, or Somewhere in Between?, *Decision Support Systems, 52*, 199-206.

Senge, P. (1992). *The Fifth Discipline: The Art and Practice of the Learning Organization*. London: Random House.

Sheremata, W. A. (2000). Centrifugal and Centripetal Forces in Radical New Product Development Under Time Pressure, *Academy of Management Review, 25*, 389–408.

Shi, X., Lu, L., Zhang, W., & Zhang, Q. (2020). Structural network embeddedness and firm incremental innovation capability: the moderating role of technology cluster. *Journal of Business & Industrial Marketing*, *36*(11), 1998–2000.

Shin, K., Kim, E. and Jeong, E. (2018). Structural Relationship and Influence between OI Capacities and Performance, *Sustainability, 10*, 1–18.

Simard, C. and West, J. (2006). Knowledge Networks and the Geographic Locus of Innovation, in *OI: Researching a New Paradigm*. Oxford: Oxford University Press (220–240).

Simsek, Z. (2009). Organizational Ambidexterity: Towards a Multilevel Understanding, *Journal of Management Studies, 46*(4), 597-624.

Sirmon, D. G., Hitt, M. A. and Ireland, R. D. (2007). Managing Firm Resources in Dynamic Environments to Create Value: Looking Inside the Black Box, *Academy of Management Review, 32,* 273–292.

Slater, S. F. and Narver, J. C. (1998). Customer-led and Market-oriented: Let's not Confuse the Two, *Strategic Management Journal, 19*(10), 1001–1006.

Slater, S. F. and Narver, J. C. (1995). Market Orientation and the Learning Organization, *Journal of Marketing, 59*(3), 63–74.

Slater, S. F. and Narver, J. C. (1999). Market-oriented is More than Being Customer-led, *Strategic Innovation Management, 20(*12), 1165–1168.

Smith, A. (1776). *The Wealth of Nations, Books I-III, with an Introduction by A. Skinner. L*. London: Penguin Books.

Smith, G., & Akram, A. (2017, December 13). Outbound OI in the Public Sector: The Roles of Intermediaries. *4th Annual World OI Conference (WOIC).*

Smith, G. and Akram, A. (2017). Outbound OI in the Public Sector: The Roles of Intermediaries', *Proceedings of the 4th Annual World OI Conference (WOIC)*, December 13-17. San Francisco.

Smith, W. K., Binns, A. and Tushman, M. (2010). Complex Business Models: Managing Strategic Paradoxes Simultaneously, *Long Range Planning, 43*, 448–461.

Smith, W. K. and Tushman, M. (2005). Managing Strategic Contradictions: A Top Management Model for Managing Innovation Streams, Organization Science, 16(5), 522–536.

Somers, M. (1994). The narrative constitution of identity: A relational and network approach. *Theory and Society*, 605–650.

Song, M., & Thieme, J. (2009). The Role of Suppliers in Market Intelligence Gathering for Radical and Incremental Innovation, *Journal of Operations Management, 26*(1), 1–22.

Sood, A. and Tellis, G. J. (2005). Technological Evolution and Radical Innovation, *Journal of Marketing, 69*, 152–168.

Sørensen, E., & Torfing, J. (2012). Introduction: Collaborative Innovation in the Public Sector. *Innovation Journal*, *17*(1), 1–14.

Sorescu, A. B., Chandy, R. K. and Prabhu, J. C. (2003). Sources and Financial Consequences of Radical Innovation: Insights from Pharmaceuticals, *Journal of Marketing, 67*(4), 82–102.

Spohrer, J. and Maglio, P. P. (2008). The Emergence of Service Science: Toward Systematic Service Innovations to Accelerate Co-Creation of Value, *Production and Operations Management, 17(*3), 238–246.

Spohrer, J. et al. (2007). Steps Toward a Science of Service Systems, *IEEE Computer, 40*(1), 71–77.

Spohrer, J. et al. (2007). Steps Toward a Science of Service Systems, *Computer, 40*, 71–77.

Stallman, R. (2002). *Free as in Freedom: Richard Stallman's Crusade for Free Software*. Boston, MA: The Free Software Foundation.

Stallman, R. (1999). Open Sources: Voices from the Open Source Revolution, in DiBona, C., Ockman, S. and Stone, M. (Ed.) *Open Sources: Voices from the Open Source Revolution*. Sebastopol, CA: O'Reilly (53–70).

Stanko, M. A., Fisher, G. J., & Bogers, M. (2017). Under the Wide Umbrella of OI. *Journal of Product Innovation Management*, *34*(4), 543–558.

Stark, J. (2002). Peer Reviews as a Quality Management Technique in Open-Source Software Development Projects, in *Quality Connection — 7th European Conference on Software Quality Helsinki*. Helsinki, Finland.

Steen, J., MacAulay, S. and Kastelle, T. (2010). A Review and Critique of the small Worlds hypothesis: The Best Network Structure for Innovation, in *Opening Up Innovation: Strategy, Organization and Technology*. Imperial College London Business School.

Stewart, K. J. and Gosain, S. (2006). The Impact of Ideology on Effectiveness in Open Source Software Development Teams, *MIS Quarterly, 30*(2), 291–314.

Suarez, F. F. and Utterback, W. Jm. (1995). Dominant Designs and the Survival of Firms, *Strategic Management Journal, 16*, 415–430.

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

Sullivan, P. H. and Fox, S. P. (1996). Establishing an Out-Licensing Activity, in *Technology Licensing: Corporate Strategies for Maximizing Value*. New York, NY: Oxford University Press (83–96).

Talke, K. (2007). Corporate Mindset of Innovating Firms: Influences on New Product Performance, *Journal of Engineering and Technology Management, 24*(1–2), 76–91.

Tambe, P. (2014). Big Data Investment, Skills, and Firm Value. *Management Science*, *60*(6), iv–vii.

Tan, F. B., & Hunter, M. G. (2003). Using Narrative Inquiry in a Study of ISProfessionals. *Proceedings of Hawaii International Conference on Systems Sciences*.

Tao, J. and Magnotta, V. (2007). Open for Business: The Air Products and Chemicals Story, in Kahn, E. (Ed.) *Innovate or Perish: Managing the Enduring Technology Company in the Global Market*. Hoboken, NJ: Wiley.

Tao, J. and Magnotta, V. (2006). How Air Products and Chemicals Identifies and Accelerates, *Research Technology Management, 49*(5), 12–18.

Tapscott, D. and Williams, A. D. (2006). *Wikinomics: How Mass Collaboration Changes Everything*. New York: Portfolio.

Tarba, S. Y., Jansen, J. J. P., Mom, T. J. M., & Raisch, S. Lawton, T. C. (2020). A Microfoundational Perspective of Organizational Ambidexterity: Critical Review and Research Directions. *Long Range Planning*, *53*(6).

Targio Hashem, I. A., Yaqoob, I., Badrul Anuar, N., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The Rise of “Big Data” on Cloud Computing: Review and Open Research Issues. *IS*, *47*, 98–115.

Taylor, A. and Greve, H. R. (2006). Superman or the Fantastic Four? Knowledge Combination and Experience in Innovative Teams, *Academy of Management Journal, 49*(49), 723–740.

Teece and Pisano, G., D. (1994). The dynamic capabilities of firms: An introduction, *Industrial and Corporate Change, 3*, 537–556.

Teece, D. J. (2015). Dynamic Capabilities and Entrepreneurial Management in Large Organizations: Toward a Theory of the (Entrepreneurial) Firm. *European Economic Review*, *86*, 202–216.

Teece, D. J. (2012). Dynamic Capabilities: Routines versus Entrepreneurial Action. *Journal of Management Studies*, *49*(8), 1395–1401.

Teece, D. J. (2010). Business Models, Business Strategy and Innovation, *Long Range Planning, 43*, 172–194.

Teece, D. J. (2007). Explicating Dynamic Capabilities: The Nature and Micro Foundations of (Sustainable) Enterprise Performance. *Strategic Management Journal*, *28*, 1319–1350.

Teece, D. J. (1998). Capturing Value from Knowledge Assets: the New Economy, Markets for Know-How, and Intangible Assets, *California Management Review, 40*, 55–79.

Teece, D. J. (2000). *Managing Intellectual Capital: Organizational, Strategic, and Policy Dimensions*. New York: Oxford University Press.

Teece, D. J. (1986). Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy, *Research Policy, 15*, 285–305.

Thompson, J. D. (1967). *Organization in Action: Social Science Bases of Administrative Theory*. New Jersey: Transaction Publisher.

Thompson, V. A. (1965). Bureaucracy and Innovation, *Administrative Science Quarterly, 10*, 1–20.

Tidd, J. (2001). Innovation Management in context: Environment, Organization and Performance, *International Journal of Management Reviews, 3*(3), 169–183.

Tien, J. M. and Berg, D. (2006). On Services Research and Education, *Journal of System Science and System Engineering, 15*(3), 257–283.

Tillson, D., Lyytinen, K. and Sørensen, C. (2010). Digital Infrastructures: The Missing IS Research Agenda, *ISResearch, 20*, 748–759.

Tilton, J. E. (1971). *International Diffusion of Technology: The Case of Semiconductors*. Washington, DC: The Brookings Institute.

Todorova, G. and Durisin, B. (2007). Absorptive Capacity: Valuing a Reconceptualization, *Academy of Management, 32*(3).

Tranekjer, T. L., & Knudsen, M. P. (2012). The (Unknown) Providers to Other firms' new product Development: What's in it for Them? . *Journal of Product Innovation Management*, *13*(4), 715–736.

Trott, P. (2005). *Innovation Management and New Product Development*. Prentice-H. Harlow.

Trott, P. and Hartman, D. (2009). Why “OI” is Old Wine in New Bottles, *International Journal of Innovation Management, 13*(4), 715–736.

Tsai, C. T., & Liao, W. F. (2011). A Study on the Framework and Indicators for OI Performance via AHP Approach. *Technology Management in the Energy Smart World*, *July-August*, 1–7.

Tschirky, H. et al. (2000). Technology Marketing: A New Core Competence of Technology: Intensive Enterprises, *International Journal of Technology Management, 20*, 459–474.

Tushman, M. and Anderson, P. (1986). Technological Discontinuities and Organizational Environments, *Administrative Science Quarterly, 31*(3), 439–465.

Tushman, M. and Murmann, J. P. (1998). Dominant Designs, Technology Cycles, and Organization Outcomes, *Academy of Management Proceedings ‘98.*

Tushman, M. and Nadler, D. (1986). Organizing for Innovation, *California Management Review, 28*(3), 74–92.

Tushman, M. and O’Reilly, C. (1996). Ambidextrous organizations: Managing Evolutionary and Revolutionary Change, *California Management Review, 38*(4), 8–30.

Tushman, M., Smith, W. K. (2002). Organizational Technology, in Baum, J. (Ed.) *Companion to Organizations.* Malden, MA: Blackwell (386–414).

Tushman, M., Smith, W. K., Wood, R. C., Westerman, G. and O'Reilly, C. (2010) Organizational Design and Innovation Streams, *Industrial and Corporate Change, 19*(5), 1331–1366.

Tushman, M. L. and Anderson, P. (1986). Technological Discontinuities and Organizational Environments, *Administrative Science Quarterly, 31*(3), 439.

Ulrich, K. (1995). The Role of Product Architecture in the Manufacturing Firm, *Research Policy, 24*, 419–440.

Un, C. A. (2010). An empirical multi-level analysis for achieving balance between incremental and radical innovations. *Journal of Engineering and Technology Management*, *27*(1–2), 1–19.

upGrad. (2021). *4 Popular Big Data Jobs – Roles & Responsibilities*. Big Data Jobs.

Usman, M. and Vanhaverbeke, W. (2017). How Start-Ups Successfully Organize and Manage OI with Large Companies, *European Journal of Innovation, 20*(1), 171–186.

Usman, M., Roijakkers, N., Vanhaverbeke, W., & Frattini, F. (2018). A Systematic Review of the Literature on Open Innovation in SMEs. In W. Vanhaverbeke, F. Frattini, N. Roijakkers, & M. Usman (Eds.), *Researching open innovation in SMEs*. World Scientific Publishing Co.

Utterback, J. (1994). *Mastering the Dynamics of Innovation*. Boston, MA: Harvard Business School Press.

Utterback, J. and Abernathy, W. J. (1975). A Dynamic Model of Process and Product Innovation, *Omega, 33*, 639–656.

Uzzi, B. and Spiro, J. (2005). Collaboration and Creativity: The Small World Problem, *American Journal of Sociology, 111*(2), 447–502.

Vaast, E. (2004). O Brother, Where are Thou?: From Communities to Networks of Practice Through Intranet Use, *Management Communication Quarterly, 18*(1), 5–44.

van de Kaa, G., & Kamp, L. (2021). Exploring design dominance in early stages of the dominance process: The case of airborne wind energy. *Journal of Cleaner Production*, *321*, 1–7.

Van de Ven, A. (1986). Central Problems in the Management of Innovation, *Manage Science, 32*(5), 590–607.

Van de Vrande, V., De Jong, J. P., Vanhaverbeke, W., & De Rochemont, M. (2009). Open Innovation in SMEs: Trends, Motives and Management Challenges. *Technovation*, *29*(6–7), 457–468.

van de Vrande, V. et al. (2009). OI in SMEs: Trends, Motives and Management Challenges, *Technovation, 29*, 423–437.

Van den Ven, A. and Garud, R. (1993). The Coevolution of Technical and Institutional Events in the Development of an Innovation, in Baum, J. and Singh, J. (Eds.) *Evolutionary Dynamics of Organizations*. Oxford: Oxford University Press.

van der Meer, H. (2007). OI – The Dutch Treat: Challenges in Thinking in Business Models, *Creativity and Innovation Management, 16*(2), 192–202.

Vandermerwe, S. and Rada, J. (1989). Servitization of Business: Adding Value by Adding Services, *European Management Journal, 6*, 314–324.

van Lieshout, J. W. F. C., van der Velden, J. M., Blomme, R. J., & Peters, P. (2021). The Interrelatedness of Organizational Ambidexterity, Dynamic Capabilities and OI: A Conceptual Model Towards a Competitive Advantage. *European Journal of Management Studies*.

Vanhaverbeke, W., & Cloodt, M. (2014). Theories of the Firms and OI. In H. Chesbrough, W. Vanhaverbeke, & J. West (Eds.), *New frontiers in OI*. Oxford University Press.

Vanhaverbecke, W., Van de Vrade, V. and Chesbrough, H. W. (2008). Understanding the Advantages of OI Practices in Corporate Venturing in Terms of Real Options, *Creativity and Innovation Management, 17*(4), 295–306.

Vanhaverbeke, W. (2006). The Inter-Organizational Context of OI, in Chesbrough, H. W., Vanhaverbeke, W., and West, J. (Eds.) *OI: Researching a New Paradigm*. Oxford: Oxford University Press (205–219).

Vanhaverbeke, W., van de Vrande, V. and Chesbrough, H. (2014). Surfing the New Wave of OI Research, in Chesbrough, H., Vanhaverbeke, W., and West, J. (Eds.) *New Frontiers in OI*. Oxford: Oxford University Press (251–258).

Varadarajan, R. (2018). Innovation and Strategy. *Review of Marketing Research*, *15*, 143–166.

Vargo, S. L. and Lusch, R. F. (2008). From Goods to Service(s): Divergences and Convergences of Logics, *Industrial Marketing Management, 37*(3), 254–259.

Vargo, S. L. and Lusch, R. F. (2006). Service-Dominant Logic: What it is, What it is not, What it might be, in Lusch, R. F. and Vargo, S. L. (Eds.) *Service-Dominant Logic of Marketing: Dialog, Debate, and Directions*. London, UK: M.E. Sharpe, Inc (43–56).

Vargo, S. L. and Lusch, R. F. (2010). From Repeat Patronage to Value Co-creation in Service Ecosystems: A Transcending Conceptualization of Relationship, *Journal Business and Marketing Management, 4*, 169–179.

Veblen, T. (1994). *The Theory of the Leisure Class: An Economic Study of Institutions*. Mineola, New York: Dover Publications Inc.

Veit, D. et al. (2014). Business Models: An ISResearch Agenda, *Business & ISEngineering, 6*(1), 45–53.

Verreynne, M.-L., de Oliveira, T., Steen, R., Indulska, M., & Ford, J. A. (2020). What Motivates 'Free' Revealing? Measuring Outbound Non-Pecuniary Openness, Innovation Types and Expectations of Future Profit Growth. *Scientometrics*, *124*, 271–301.

Veryzer, R. W. j. (1998). Discontinuous Innovation and the New Product Development Process, *Journal of Product Innovation Management, 15*, 304–321.

Veugelers, R., & Cassiman, B. (2006). In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition. *Management Science*, *52*(1), iv–154.

Ville, S. (2011). Historical Approaches to Creativity and Innovation, in Mann, L. and Chan, J. (Eds.) *Creativity and Innovation in Business and Beyond: Social Science Perspectives and Policy Implications*. New York (64–81).

Ville, S. (1990). *Transport and the Development of the European Economy, 1789-1914*. Houndsmills: Macmillan.

Ville, S. (2011). Historical Approaches to Creativity and Innovation, in Mann, L. and Chan, J. (Eds.) *Creativity and Innovation in Business and Beyond: Social Science Perspectives and Policy Implications*. New York: Routledge (64–81).

Vivieros Lopes, A. P. V. B., & Monteiro de Carvalho, M. (2018). Evolution of the OI Paradigm: Towards a Contingent Conceptual Model. *Technological Forcasting and Social Change*, *132*, 284–298.

von Hippel, E. (1986). Lead Users: A Source of Novel Product Concepts, M*anagement Science, 32*(7), 791–805.

von Hippel, E. (2001). Open Source Shows the Way: Innovation by and for Users—No Manufacturer Required, *Sloan Management Review, 42*(4), 82–86.

von Hippel, E. (1988). *The Sources of Innovation*. New York, NY: Oxford University Press.

von Hippel, E. (2005). *Democratizing Innovation*. Cambridge, MA: MIT Press.

von Hippel, E. and Katz, R. (2002). Shifting Innovation to Users via Toolkits, *Management Science, 48*, 821–834.

von Krogh, G., Nonaka, I. and Rechsteiner, L. (2012). Leadership in Organizational Knowledge Creation: A Review and Framework, *Journal of Management Studies, 49*(1), 240–277.

von Krogh, G. and von Hippel, E. (2006). The Promise of Research on Open Source, *Management Science, 52*(7), 975–983.

Vontris, D. et al. (2017). Ambidexterity, External Knowledge and Performance in Knowledge-Intensive Firms, *Journal of Technology Transfer, 42*, 374–388.

Wade, J. (1995). Dynamics of Organizational Communities and Technological Bandwagons: An Empirical Investigation of Community Evolution in the Microprocessor Market, *Strategic Management Journal, 16*, 111–134.

Wadhwa, A. and Kotha, S. (2006). Knowledge Creation Through External Venturing: Evidence from the Telecommunications Equipment Manufacturing Industry, *Academy of Management Journal, 49*, 819–835.

Wagner, S. M. (2012). Tapping Supplier Innovation, *Journal of Supply Chain Management, 48*(2), 37–52.

Wagner, S. M., & Bode, M. (2014). Supplier relationship-specific investments and the role of safeguards for supplier innovation sharing. *Journal of Operations Management*, *32*, 65–78.

Walsham, G. (2012). Are We Making a Better World with ICTs? Reflections on a Future A genda for the IS Field, *Journal of Information Technology, 27*(2), 87–93.

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

Walsham, G. (1993). *Interpreting ISin Organizations*. Chichester: Wiley.

Walsham, G. (1995). Learning about Being Critical, *ISJournal, 15*(2), 111–117.

Wang, Y., & Rajagopalan, N. (2015). Alliance Capabilities: Review and Research Agenda. *Journal of Management*, *41*, 236–260.

Wang, Z. and Wang, N. (2012). Knowledge sharing, innovation and firm performance, *Expert Systems with Applications, 39*(10), 8899–8908.

Ward-Schofield, J. (1993). Increasing the Generalizability of Quantitative Research. In *Social Research, Philosophy, Politics and Practice* (Hammersley). Sage.

Watson, R. T. et al. (2008). The Business of Open Source: Tracking the Changing Competitive Conditions of the Software Industry, *Communications ACM, 51*(4), 41–46.

Webb, B., & Mallon, B. (2007). A Method to Bridge the Gap between Breadth and Depth in IS Narrative Analysis. *JAIS*, *8*(7), 369–382.

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

Webster, F. E. J. (1992). The Changing Role of Marketing in the Corporation, *Journal of Marketing, 56*, 1–17.

Wei, Z., Yi, Y. and Guo, H. (2014). Organizational Learning Ambidexterity, Strategic Flexibility, and New Product Development, *Journal of Product Innovation Management, 31*(4), 832–847.

Weick, K. E. (1979). *The Social Psychology of Organizing*. New York: McGraw-Hill.

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

Wenger, E. (1998). Communities of Practice: Learning as a Social System, *Systems Thinker, June*.

Wenger, E., McDermott, R. and Snyder, W. (2002). *A Guide to Managing Knowledge: Cultivating Communities of Practice*. Boston, Massachusetts: Harvard Business School Press.

Wermelinger, M. and Yu, Y. (2008). Graph-centric tools for understanding the evolution and relationships of software structures, *Proceedings of the 15th Working Conference on Reverse Engineering*. Antwerp, Belgium (329).

West, J. (2003). How Open is Open Enough? Melding Proprietary and Open Source Platform Strategies, *Research Policy, 32*, 1259–1285.

West, J. and Gallagher, S. (2006). Challenges of OI: the Paradox of Firm Investment in Open-Source Software, *R&D Management, 36*(3), 319–331.

West, J. and Lakhani, K. R. (2008). Getting Clear About Communities in OI, *Industry and Innovation, 15*(2), 223–231.

West, J. and Mace, M. (2010). Browsing as the Killer app: Explaining the Rapid Success of Apple's iPhone, *Telecommunications Policy, 34*, 270–286.

West, J. and O'Mahony, S. (2005). Contrasting Community Building in Sponsored and Community Founded Open Source Projects, in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*. Waikoloa, Hawaii.

West, J., Vanhaverbeke, W. and Chesbrough, H. W. (2006). OI: A Research Agenda, in *OI: Researching a New Paradigm*. Oxford: Oxford University Press (285–307).

West, M. A. and Anderson, N. R. (1996). Innovation in Top Management Teams, *Journal of Applied Psychology, 81*, 680–693.

Wilkes, L. (2006). *Applying ESB, Best Practice Report*, CBDI Report, 23–31.

Williiams, M. (2000). Interpretivism and Generalisation. *Sociology*, *34*(2), 209–224.

Windrum, P., & Koch, P. M. (n.d.). *Innovation in Public Sector Services: Entrepreneurship, Creativity and Management*. Edward Elgar Publishing.

Wirtz, B. W. (2011). *Business Model Management. Design – Instruments – Success Factors*. Gabler Verlag.

Wirtz, B. W., Schilke, O. and Ullrich, S. (2010). Strategic Development of Business Models: Implications of the Web 2.0 for Creating Value on the Internet, *Long Range Planning, 43*, 272–290.

Witell, L. et al. (2016). Defining Service Innovation: A Review and Synthesis, *Journal of Business Research, 69*, 2863–2872.

Witzeman, S. et al. (2006). Harnessing External Technology for Innovation, *Research-Technology Management, 49*(3), 19–27.

Wladawsky-Berger, I. (2006). People-Oriented, Services-Intensive, Market-Facing Complex Systems, in *IBM Service Science Management and Engineering Conference*: Education for the 21st. Palisades, NY.

Wolf, V. (2019). Ambidexterity in Service Innovation Research: A Systematic Literature Review', in *14th International Conference on Wirtschaftsinformatik*. Siegen, Germany.

Wong, A., Tjosvold, D. and Liu, C. (2008). Innovation by Teams in Shanghai, China: Cooperative Goals for Group Confidence and Persistence, *British Journal of Management, 20*(2), 238–251.

Woolridge, R. W., McManus, D. J. and Hale, J. E. (2007). *Stakeholder' Management Decision, 57*(3), 703-723.

Wozniak, G. (1993). Joint Information Acquisition and new Technology Adoption: Late versus Early Adoption. *The Review of Economics and Statistics*, *75*(3), 438–445.

Xie, Y., Xu, K., & Huang, J. Q. (2020). How Do Innovation Intermediaries Influence Outbound OI in China? A Moderated Mediation Mechanism. *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 344–348.

Yaqoob, I., Targio Hashem, I. A., Gani, A., Mokhtar, S., Ahmed, E., Badrul Anuar, N., & Vasilakos, A. v. (2016). Big Data: From Beginning to Future. *International Journal of Information Management*, *36*(6B), 1231–1247.

Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods*. Thousand Oak, California: Sage Publications, Inc.

Yin, R. K. (2003). *Case Study Research, Design and Methods*. Thousand Oak, California: Sage Publications, Inc.

Yoo, S. (2011). *Development of an Instrument to Measure Desorptive Capacity*. Sungkyunkwan University, Suwon, South Korea.

Yoo, Y. et al. (2012). Organizing for Innovation in the Digitized World, *Organization Science, 23*, 1398–1408.

Zheng, F., Jiao, H., & Cai, H. (2018). Reappraisal of Outbound OI under the Policy of China's 'Market for Technology.' *Technology Analysis & Strategic Management*, *30*(1), 1–14.

Zhoa, J. L., Tanniru, M. and Zhand, L. (2007). Services Computing as the Foundation of Enterprise Agility: Overview of Recent Advances and Introduction to the Special Issue, *Frontiers of Information System, 1*, 1–8.

Zhu, K. (2004). The Complementarity of Information Technology Infrastructure and E-Commerce Capability: A Resource-Based Assessment of their Business Value, *Journal of Management Information Systems, 21*(1), 167–202.

Ziegler, N. et al. (2013). Creating value through external intellectual property commercialization: A desorptive capacity view, *Journal of Technology Transfer, 38*(6), 930–949.

Zimmermann, R., Ferreira, L. M. D., & Moreira, A. C. (2016). The influence of supply chain on the innovation process: A systematic literature review. *Supply Chain Management: An International Journal*, *21*(3), 16.

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

Zook, C. and Rigby, D. K. (2002). Open-Market Innovation, *Harvard Business Review*, (October Issue).

Zott, C. and Amit, R. (2008). The Fit between Product Market Strategy and Business Models: Implications for Firm Performance, *Strategic Management Journal, 29*, 1–26.

Zott, C., Amit, R. and Massa, L. (2011). The Business Model: Recent Developments and Future Research, *Journal of Management, 37*(4).

Zuboff, S. (1988). *Dilemmas of Transformation in the Age of the Smart Machine: The Future of Work and Power*. New York: Basic Books.

Zuboff, S. and Maxmin, J. (2004). *The Support Economy: Why Corporations are Failing Individuals and the Next Episode of Capitalism*. New York, NY: Viking.

APENDIX A

Table A1. List of Organisations and Participants Professions

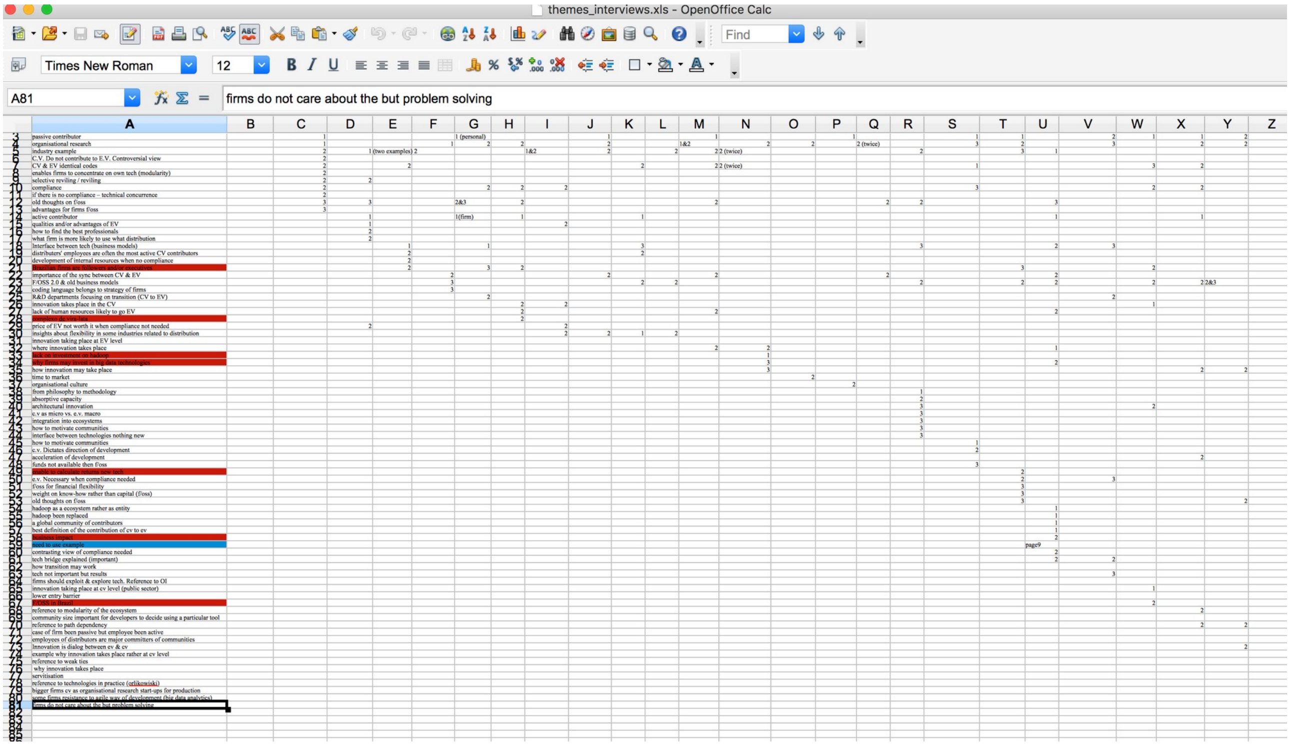
Firm	Sector	Type	Participant Profession
F1	Banking	Multinational	IT Business Solution Specialist
F2	IT Service Firm & Hardware	Multinational	Head of Analytics and Data Management
F3	Benefits and Rewards Service	Multinational	Business Analytics Manager
F4	IT Service Firm	Multinational	Lead Tecnical Specialist
F5	IT Service FirmIT Service Firm	Domestic	Java Software Engineer
F6	Telecom & IT Service Firm	Multinational	IT Project Manager
F7	IT Service Firm & IT Service Firm	Multinational	Senior Data and Software Engineer
F8	IT Service Firm	Domestic	Big Data Architect
F8	IT Service Firm	Domestic	Data Analyst
F9	Aviation	Multinational	Technical Account Manager
F10	IT Service Firm	Start-up	Innovation Advisory Consultant
F11	IT Service Firm	Domestic	Lecture & Systems Administrator
F12	Telecommunication	Multinational	(MVP) Software Engineer
F13	IT Service Firm	Domestic	Big Data Architect
F14	IT Service Firm	Domestic	Data Scientist & Project Manager
F15	Online Retail	Domestic	Data Engineer
F16	IT Service Firm (Public Sector)	Start-up	Chief Executive Officer
F17	IT Service Firm & Hardware	Multinational	Advisory Solutions Architect
F18	IT Service Firm	Multinational	Big Data Architect
F18	IT Service Firm	Multinational	Big Data & IoT Architect
F19	Banking	Domestic	Software Engineer

F20	IT Service Firm	Multinational	Analytics Sales Executive
F21	IT Service Firm	Domestic	Chief Executive Officer
F22	IT Service Firm (Public Sector)	Domestic	Chief Executive Officer
F23	IT Service Firm	Start-up	Chief Executive Officer
F24	Media	Domestic	Big Data Engineer
F25	IT Service Firm	Multinational	Data Engineer & Team Leader

APPENDIX B.

Figure B1

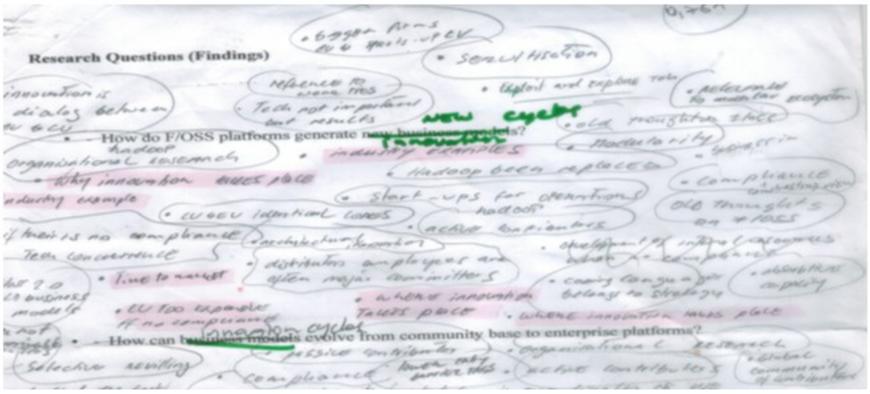
Themes and Ideas

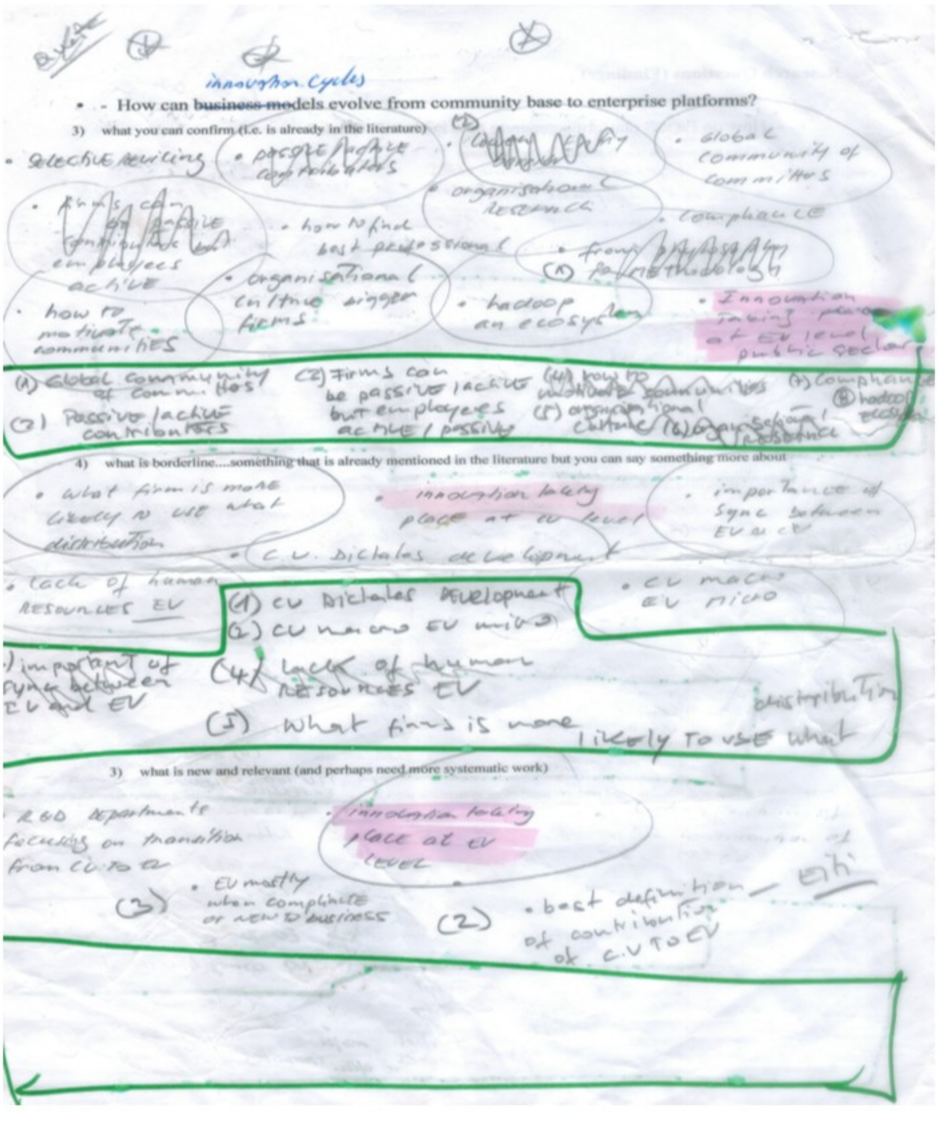


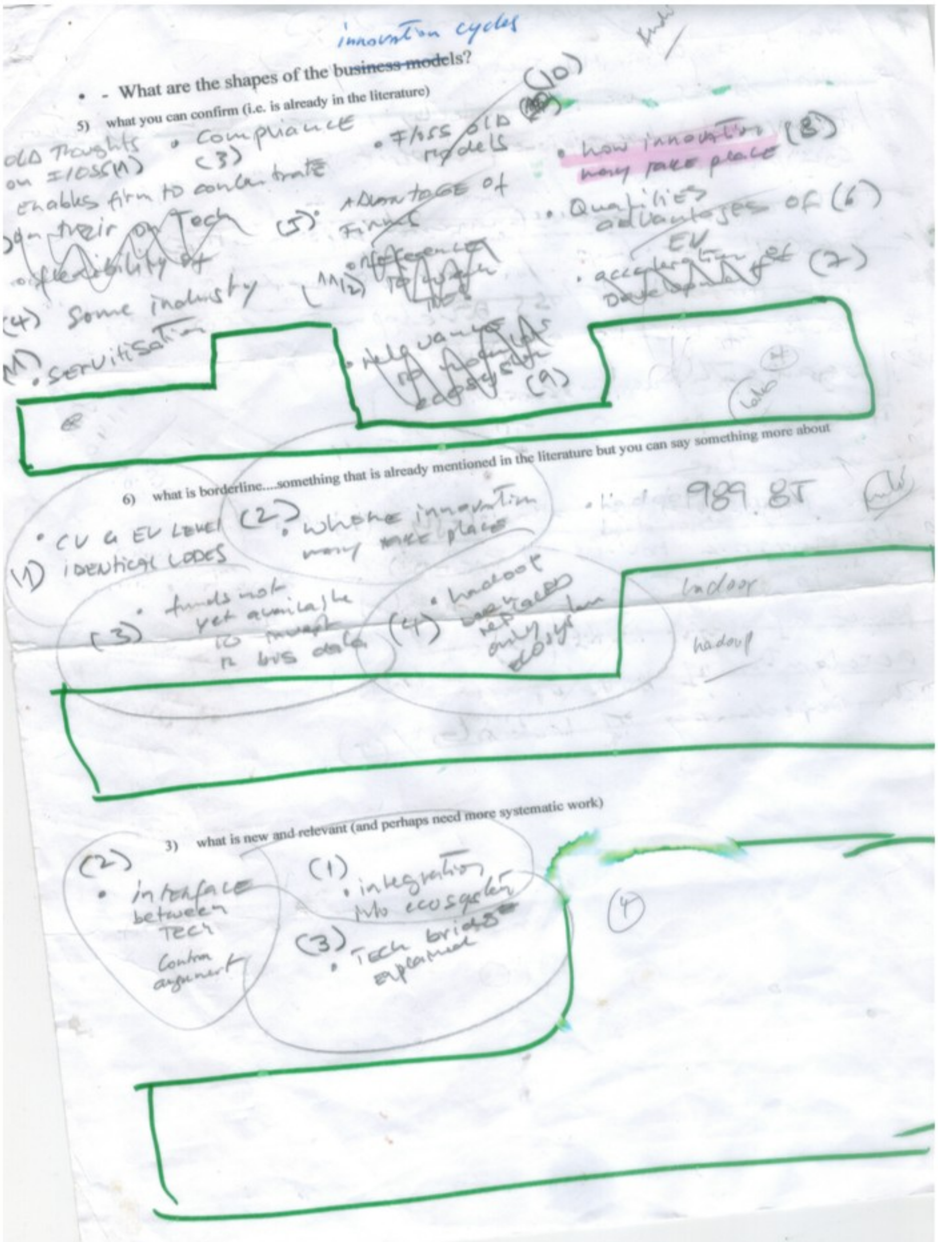
APPENDIX C

Figure C1.

Coding Documents







- - What are the shapes of the ~~business models~~?

5) what you can confirm (i.e. is already in the literature)

6) what is borderline....something that is already mentioned in the literature but you can say something more about

3) what is new and relevant (and perhaps need more systematic work)

APPENDIX D

Sub-questions

Contextualisation questions

What is your job title?

Can you explain your role within your organisation?

Can you explain your involvement within the Apache Hadoop community?

Links between C.V and E.V.

Is the C.V . contributing to the E.V?

How does the development of C.V . affect the deployment of the E.V .?

Is there a translation of ideas between the two? How does this exchange of ideas work?

What is the link between the C.V . and E.V . in your organisation?

How C.V. ideas become E.V. services and later develop in to "new" business models

What is the importance of new service developed in the "Community Version" for the commercialisation of "new" business models proposed for your current market?

Could you explain why the "Community Version" is so important for the development of "new" business models for the "Enterprise Version"?

How does the “Community Version” assist your organisation in advancing new business models for your current market (E.V.)?

How does the relationship between “Community Version” and “Enterprise Version” affect the development of your organisation’s business models for both versions?