Disruption: A Transdisciplinary User-Centric Framework for Innovation Through Data Design and Analytics

By: Mahya Mirzaei

Supervisor: Professor Mary-Anne Williams

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Faculty of Engineering and Information Technology
University of Technology Sydney, Australia

September 2017
Declaration of Originality

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as part of the collaborative doctoral degree and/or fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This research is supported by an Australian Government Research Training Program Scholarship.

Signature of Student: 

Date: 10 March 2018
First and foremost I would like to thank my PhD supervisor Professor Mary-Anne Williams who is nothing short of the most incredible and inspiring teacher that any PhD student could wish to have the priviledge of working with. Not only did Professor William’s immense technical expertise helped shape this thesis, but also more importantly, her grand vision, mentorship, and heart of gold helped me find my place in the world, for which I will forever be grateful.

I would also like to thank my mentors at IBM Research Australia, Dr. Juerg Von Kaenel (Associate Director of IBM Research Australia) and Dr. Kent Steer, for their valuable feedback and direction towards the completion of my thesis specially in regards to the data generator I created during my time at IBM.

This research would also not have been possible without the help of Gavin Smith, Bradley Austin, Vernon Young, Alex Smith, and Dorothy Green from Commonwealth Bank of Australia (CBA) who worked extremely hard to grant me permission to be part of CBA projects and conduct my research. I am also extremely grateful for the countless hours they spent providing me with valuable advice and feedback which truly helped form my thesis.

I also would like to thank Startup Catalyst, for assisting me in organising meetings with some of the most accomplished executives and managers in Silicon Valley’s most prominent organisations. Your help is appreciated immenesly.

Furthermore, I would like to thank my fellow researchers at Magic Lab UTS for their continual support and assistance. I am grateful for all the times that all of you listened to my PhD update and gave me valuable feedback.

And lastly I would like to dedicate this thesis:

To my darling husband who has been the shining light even in my darkest paths,
To my parents who have given me everything I have today, and have believed in me with the air of an unearthly faith,

And lastly to my sisters whose little hearts I am forever grateful for.
3.3.2.1 Disruptive innovation .......................... 33
3.3.2.2 Big-bang disruption ........................... 35

3.4 Innovation success factors .......................... 39
  3.4.1 Innovation leadership ............................ 40
  3.4.2 Managerial levers ............................... 42
    3.4.2.1 Organisational mission and strategy ............ 43
    3.4.2.2 Resource allocation ............................ 44
    3.4.2.3 Structure and systems .......................... 45
    3.4.2.4 Organisational culture ........................ 47
  3.4.3 Business processes ............................... 48
    3.4.3.1 Structured problems ........................... 52
    3.4.3.2 Unstructured problems .......................... 54
  3.4.4 Customers ....................................... 62
    3.4.4.1 Design thinking ............................... 64
    3.4.4.2 Understanding design thinking as a concept .... 67

4 Research process ........................................ 73
  4.1 Philosophical assumptions of positivism and interpretivism ........ 73
    4.1.1 Positivist perspective ............................ 74
    4.1.2 Interpretive perspective .......................... 74
  4.2 Research approaches .................................. 76
    4.2.1 Quantitative research .............................. 77
    4.2.2 Qualitative research ............................... 78
    4.2.3 Taxonomy of research approaches .................... 78
4.3 Research method chosen for this study

4.3.1 Case study method

4.3.2 Limitation of research method

4.4 Data collection

4.4.1 Case study design

4.4.2 Case participants

4.5 Data analysis

4.6 Grounded theory

4.6.1 Grounded theory procedures

4.6.1.1 Coding

4.6.1.2 Memo writing

4.6.1.3 Theoretical sampling and comparison

5 Field studies: case study analysis

5.1 Case study analysis

5.1.1 Propositions that emerged from the data analysis

5.2 Results of case study analysis

5.2.1 Importance of the problem

5.2.1.1 Understand the business challenges and problems

5.2.1.2 Understand customers and focus on them

5.2.1.3 Focus on ideas that result in 10x better

5.2.2 People aspect

5.2.2.1 Quality of people

5.2.2.2 Team work
5.2.3 Internal organisational culture ........................................ 118
  5.2.3.1 Innovation culture ............................................. 119
  5.2.3.2 Failure accepting culture ..................................... 121
  5.2.3.3 Innovation framework ......................................... 122

5.2.4 Experimentation factors ............................................. 124
  5.2.4.1 Prototyping .................................................... 125
  5.2.4.2 Iteration ...................................................... 125
  5.2.4.3 Fast-paced execution ......................................... 127

5.2.5 Managerial support .................................................. 128
  5.2.5.1 Managers as team members ................................... 129
  5.2.5.2 Adequate resources .......................................... 130

5.2.6 Use variety of data .................................................. 130

5.2.7 Organisational structure ............................................ 132
  5.2.7.1 Physical environment ......................................... 133
  5.2.7.2 Separate small units ......................................... 133

5.3 Framework for innovative analytics ................................ 134
  5.3.1 The prerequisite phase .......................................... 140
    5.3.1.1 Managerial support ....................................... 140
    5.3.1.2 Innovation culture ....................................... 143
    5.3.1.3 People aspect ............................................. 144
    5.3.1.4 Physical environment .................................... 145

5.3.2 User-centric data analytics phase ................................. 148
    5.3.2.1 Empathy and framing .................................... 152
Appendix 1
  Interview questions for Australia’s largest financial institutions  221

Appendix 2
  Interview questions for large innovative organisations in Silicon
  Valley                                                        225

Appendix 3
  Interview questions for managers involved in the CDA projects at
  Commonwealth Bank of Australia                                 229

Appendix 4
  Data design process for the transactional data generator        231

References                                                    241
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Definitions of big data based on an online survey of 154 global executives in April 2012.</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>(adapted from Gandomi and Haider (2015))</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>Processes for extracting insights from big data (Labrinidis and Jagadish, 2012)</td>
<td>16</td>
</tr>
<tr>
<td>2.3</td>
<td>Phases of the CRISP-DM reference model</td>
<td>22</td>
</tr>
<tr>
<td>3.1</td>
<td>Methodology for defining the strategic significance of disruptive innovations (Christensen, 1995)</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>The phases of big bang disruption (Downes and Dunning, 2014)</td>
<td>38</td>
</tr>
<tr>
<td>3.3</td>
<td>Innovation Type Matrix (Verworn et al., 2007)</td>
<td>52</td>
</tr>
<tr>
<td>3.4</td>
<td>Stage-gate process of the early phase (Cooper and Kleinschmidt, 1988)</td>
<td>54</td>
</tr>
<tr>
<td>3.5</td>
<td>Example of GE’s probe and learn process (Lynn et al., 1996)</td>
<td>56</td>
</tr>
<tr>
<td>3.6</td>
<td>The NCD model (Koen et al., 2002)</td>
<td>58</td>
</tr>
<tr>
<td>3.7</td>
<td>The magic ideation quadrant diagram (Cooper et al., 2008)</td>
<td>65</td>
</tr>
<tr>
<td>3.8</td>
<td>Description of a DT process (source Stanford d.School, 2009)</td>
<td>70</td>
</tr>
<tr>
<td>5.1</td>
<td>The creative data analytics framework (CDA)</td>
<td>137</td>
</tr>
<tr>
<td>6.1</td>
<td>Business model canvas for general insurance completed with the leadership team (co-created with CBA)</td>
<td>164</td>
</tr>
<tr>
<td>6.2</td>
<td>Challenge statements (co-created with CBA)</td>
<td>166</td>
</tr>
</tbody>
</table>
6.23 Grouping the concepts (co-created with CBA) ............... 185
6.24 Sample prototype - Offering retension incentives (co-created with
CBA) ................................................................. 186
6.25 Sample prototype - Setting up a new policy (co-created with CBA) . 187
6.26 Sample prototype - Moving to new address for under 30s (co-created
with CBA) ................................................................. 187
6.27 Sample prototype - Insure your child’s toy (co-created with CBA) . . 188
6.28 Sample prototype - Customisable renters product (co-created with
CBA) ................................................................. 188
6.29 Sample prototype - Content insurance (co-created with CBA) ....... 189
6.30 Sample prototype - Renewal with competitor (co-created with CBA) 189
6.31 Sample prototype - Renovation poke (co-created with CBA) ......... 190
6.32 Sample prototype - Insure the device for free (co-created with CBA) 190
6.33 Sample prototype - No excess (co-created with CBA) ............. 191
6.34 Sample prototype - Commit to Fix (co-created with CBA) ........... 191
6.35 Sample prototype - Empty investment properties (co-created with
CBA) ................................................................. 192
6.36 Paths that prototypes could take based on testing results (co-created
with CBA) ................................................................. 193
6.37 Tested prototype - Setting up a new policy (co-created with CBA) . 194
6.38 Tested prototype - Moving to new address for under 30s (co-created
with CBA) ................................................................. 194
6.39 Tested prototype - Leaving customers (co-created with CBA) ....... 195
6.40 Tested prototype - Insure your child’s toy (co-created with CBA) . . 195
6.41 Tested prototype - Customisable renters product (co-created with
CBA) ................................................................. 196
6.42 Tested prototype - Content insurance (co-created with CBA) ... 196
6.43 Tested prototype - Renewal with a competitor (co-created with CBA) 197
6.44 Tested prototype - Renovation poke (co-created with CBA) ... 197
6.45 Tested prototype - Insuring the device for free (co-created with CBA) 198
6.46 Tested prototype - No excess (co-created with CBA) ........ 198
6.47 Tested prototype - Commit to fix (co-created with CBA) ... 199
6.48 Tested prototype - Empty investment properties (co-created with CBA) .......................................................... 199
6.49 A summary of the achievements of the project (co-created with CBA) 200
6.50 A typical data architecture at Commwealth Bank (co-created with CBA) .......................................................... 201
List of Tables

2.1 List of organisations involved in our qualitative research ...... 17
3.1 Theories used in highly cited papers (Crossan and Apaydin, 2010) . 28
3.2 Fives generations of innovation models (Adopted from Rothwell (1994) and Xu et al. (2007)). ........................................ 50
3.3 Studies confirming the impact of the front end on NPD ........... 51
4.1 Information systems research approaches (Galliers, 1992) ....... 80
5.1 List of organisations that participated in our case studies ........ 104
5.2 Core categories and subcategories that emerged from the data analysis.107
5.3 Propositions that emerged from data analysis .................... 109
5.4 Entrepreneurial and innovation culture factors (Drucker, 1993) . . . 120
A 1.1 Transformation questions ........................................... 223
A 2.1 Transformation questions ........................................... 227
A 4.1 Age group and mean income of customer categories ............ 234
Abstract

While innovation has always been critical for competitiveness of businesses, fierce competition resulting from the global economy and constant waves of disruption has made innovation even more crucial for the survival of large organisations. Today, extremely large volumes of data from variety of sources are continuously created with immense speed. Containing deep information and insights into customer habits and needs, data has the potential to become a key enabler of competition and innovation. In the financial industry in particular, with no physical products, data is the most valuable asset that needs to be utilised to create competitive advantage and innovation.

However, studies in literature as well as primary qualitative research that I have conducted in collaboration with Ernst and Young, reveal that financial institutions are falling short of exploiting data and analytics' full potential for innovation and competition. This is due to the failure to discover high-value problems that may be solved using data analytics, which can have the potential to result in significant value for customers and business.

The aim of this research is to develop and evaluate a holistic model that increases the probability of success of data analytics endeavours in large organisations, resulting in high value and innovative products and services. Using the research methodologies of interpretive case studies and grounded theory for data analysis, I derived the influencing factors for the success of developing innovation using data analytics and hence created a framework that mapped these success factors in a cohesive and clear way. The generated framework, referred to as Creative Data Analytics (CDA), provides a methodology that consists of both creative and analytical techniques which enable organisations to develop an end to end roadmap to creative data analytics innovations. The CDA Framework integrates customer needs and predictive data analytics, and directs the investigation of data towards an innovative solution with a higher probability of solving an important real customer need or business problem.
The validity of the CDA framework was evaluated by conducting action research using three projects involving data analytics at the Commonwealth Bank of Australia. These projects were conducted according to the CDA framework principles and the degree of innovation of the solutions derived from these projects were evaluated qualitatively by interviewing managers and innovation experts involved in the projects.
1 Introduction

This chapter presents an overview of the ideas and issues that form the basis of this thesis. Many of the issues presented in this chapter will be reviewed in depth in later chapters, however, the aim of this introduction is to provide a brief overview of the key ideas and constructs forming this thesis.

The primary concern for this study is the effective use of data analytics for innovation in large financial organisations. Data analytics has the potential to offer unprecedented competitive advantage and innovation, and the finance sector - faced with immense competition and disruption - are desirous to unlock the power of data. While data analytics is believed to offer immense potential for maintaining competitiveness, the use of data analytics for innovation is vastly under-examined, with very few empirical publications linking specific activities in analytics to the outcomes of analytics projects or overall innovation success.

This study focuses on how the financial organisations can effectively take advantage of the massive amounts of data they possess (or can obtain) to create innovative products and services that generate significant business or customer value. The purpose of this research is to generate a grounded theory as opposed to testing a theory that has been determined a priori. A grounded theory is a theory that is discovered, developed, and provisionally verified through systematic data collection and analysis of data relating to a specific phenomenon (Strauss and Corbin, 1990). It aims to build on the work done so far on cultivating innovation in large organisations, by using the grounded theory method to develop a theory designed to increase the understanding of the factors that influence the effective use of data analytics for innovation. The success factors developed in the theory are then used to create a framework for conducting data analytics in a way that increases the probability of success of creating innovative products and services that generate crucial business and customer value.

Many research in Information Systems (IS) have been characterised by the application of traditional quantitative methodologies, therefore, experiments and descrip-
tive questionnaire surveys in large-N studies have been the norm. In the IS field, however, there are increasing calls for more qualitative research. Markus (1987) stated that IS comprises of a wide variety of topics that are best explored by means of qualitative methods of inquiry. One such topic, I believe, is investigation into the deep understanding of the impact of data analytics for innovation in organisations. There have been relatively few published qualitative research studies on the theme of data analytics usage in organisations and most of the research to date has been technical or descriptive in nature. This is despite the prominent impact of qualitative research on IS research more generally. Research into the non-technical aspects of the implementation of data analytics in large organisations to gain competitive advantage and business and customer value, has largely been ignored in most available literature.

The majority of research to date addresses the data storage, management, and processing tools, analytics tools and methods, and visualization and evaluation tools. Limited research has been conducted into the ways analytics can be adopted by organisations to drive value and the actions required to increase the probability of success of such attempts. Furthermore, to date, there has been no established theory in the literature for the use of data analytics for innovation in large organisations. While models of data analytics use for value creation exist (e.g. Stubbs, 2014; Bernard, 2015) these fail to provide a holistic and end-to-end framework that guides organisations in a practical sense, in ways that they can increase the probability of deriving value and successful innovation from analytics.

The data used in most studies in this field have been collected via semi-structured interviews or surveys and used for descriptive purposes and significantly limited theory development or testing has been reported in this area. More research of a qualitative nature is required to gain a greater understanding of the phenomenon of innovation using data analytics from the perspectives of the people involved in their natural settings. We need "in-depth, penetrating investigation that strives for relational understanding of all the various factors that comprise and affect the object of the study" (Wolf, 1979).

Literature review shows that there are many gaps in terms of knowledge and understanding of the factors that influence the success of data analytics in the creation of innovations that result in significant customer and business value. In particular there is a lack of attention to the contextual aspects of this. Hence, some of the possible unexplored areas of past research relate to the factors influencing the successful use of data analytics for innovation and value creation.
Therefore, this research will try to fill this gap by conducting an empirical study that examines the factors that influence the success of data analytics projects in creating innovative products and services that deliver significant business and/or customer value. I argue that research and practice on data analytics for value creation in organisations have reached a stage where theory development is critical to the accumulation of new knowledge. The development of theory about the use of data analytics for innovation, grounded in data gathered from large innovative enterprises, will contribute significantly to the field of IS and guides in the development of policy for the adoption of data analytics for value creation and innovation in organisations. Using the success factors derived from the theory, we will create a holistic and end-to-end framework for the use of data analytics for innovation in large organisations.

1.1 Background

Innovation, which was coined by Schumpeter in 1911, is a critical source of competitive advantage and both scholars and practitioners understand its crucial role in maintaining business performance in an increasingly changing environment (Dess and Picken, 2000; Tushman and O’Reilly, 1996).

For the finance sector, innovation and meeting customer needs have never been more crucial. Despite record profits, financial institutions are under pressure to maintain high levels of performance. New technologies and customer habits are eroding traditional revenues and the banking sector today faces the same disruptive forces that have destroyed business models in the media, publishing and retail industries. In 2014, Ian Narev the CEO of Commonwealth bank of Australia said: “we look at other industries which have undergone major technology-based upheaval and our fundamental mind-set is that this could and will happen here. The nature of the change is so fundamental that it is difficult to foresee an industry where it’s not some combination of a significant threat or a significant opportunity, or both. When the change starts to accelerate it will do it rapidly and therefore the time to feel the urgency is now.”

Today competition is no longer only coming from players within the same industry and financial institutions are faced with massive new, technology-rich non-bank competitors such as well-resourced global companies like Apple and Google that are utilising the disruptive effects of faster internet speeds, new devices, data and applications. Furthermore, emerging alternative non-bank players – more commonly known as Fintechs – are beginning to transform the financial sector by revamping offerings and solutions in the new digital landscape. Using technology at their
core, and providing customers solutions to their deepest needs, pains and wishes, they are disrupting the financial sector, chipping away at its borders. While the high regulations surrounding the finance sector is protecting the incumbent’s core business for now, it is short-sighted to believe that regulations won’t crumble under the pressure customers will exert to obtain the services they truly desire. Take the highly regulated Taxi industry as an example, whose strong regulation gave it a false sense of security, justifying its lack of innovation, believing itself to be impenetrable. However, Uber with its innovation that had customers’ deepest needs at its heart by solving their core pain point in regards to transportation, came along and under immense pressure from consumers, governments had to give-in and legalise the use of ride sharing.

Disruptive innovations and big-bang disruptions fuelled by new technologies and digitization are posing immense threats to incumbents in all sectors including the finance sector. For financial institutions, the largest risk of all may be to do nothing – ignoring the waves of disruption in business and society has led to extinction for many. Examples of this include the case of Nokia – which in spite of owning 30 per cent of the world’s mobile patents (GSM) finds itself struggling to stay relevant; Kodak invented the digital camera technology, but failed to capitalise on it; General Motors – produced the first commercial electric car almost 20 years ago, stopped its rollout, and now finds itself trying to catch up with the Japanese hybrids; bookstores, the music industry and newspapers – all of which were the giants of their time, and now find themselves trying to avoid elimination.

Consequently, in today’s era of digitization and disruptions (particularly in the forms of disruptive innovations and big-bang disruptions), innovation is more important than ever before. For financial institutions, with no physical products, data – the source of information - is the most valuable asset they possess and one that can be the enabler of unprecedented competitive advantage and innovation. Data contains valuable insights into customers’ preferences and needs, allowing it to be used as a basis for the creation of products and services that remedy emerging problems or opportunities in a timely and more competitive manner.

The power of data and the potential for it to be the source of immense innovation, has not been overlooked by financial institutions. In fact, according to our qualitative study (discussed in chapter 2.3.), in an attempt to unlock the power of data for value creation, every major financial institution in Australia has started a dedicated unit for data analysis in the past two to three years. However, according to our study that was conducted in collaboration with Ernst and Young, the success of such endeavours have been limited in terms of value creation, due to the fact that the organisations are
unable to routinely find high value problems and areas to direct their data analytics efforts to. As evidenced by our qualitative study that included conducting semi-structured interviews with 13 major financial institutions in Australia, currently the most difficult part of using analytics for competitive advantage and innovation is indeed finding and devising high value business and/or customer problems. In literature, there is currently no framework that can assist organisations with this aspect of their data analytics projects and while data analytics is considered very important in literature for maintaining competitiveness, the use of data analytics for innovation is vastly under-examined, with very few empirical publications linking specific activities in the analytics to the outcomes of analytics projects or overall innovation success.

1.2 Research goals and question

It is evident that the finance industry must innovate to stay competitive in today’s uncertain and ever changing economy. While data analytics has the potential to be a source of innovation and hence create immense competitive advantage and value for both customers and businesses, the finance sector finds it difficult to find what projects/products and services to focus their data analytics efforts on, in order to obtain significant business and customer value. We can formulate our central research question as follows:

*How to create a holistic end-to-end framework that increases the probability of creating innovative products and services using data analytics, which result in significant business or customer value?*

Furthermore, this study has three related research objectives:

Objective 1: To identify the *success factors* for creating innovative services and products using data analytics.

Objective 2: To build a *grounded theory* for factors affecting the success of data analytics projects resulting in innovative products and services.

Objective 3: To propose a *holistic framework* that provides a methodology to enable organisations to develop an end to end roadmap to creative data analytics which increases the probability of successfully creating innovative products or services.

To tackle the main research question in this thesis, I conducted multiple case studies of 12 of the most innovative large organisations in Silicon Valley that have successfully used data analytics to create innovative services and products that have resulted
in significant business or customer value. The aim is to gain a deeper insight into the issues raised by the financial institutions in our study with Ernst and Young (see section 2.3) and to understand the factors that influence the success of data analytics endeavours resulting in innovation. It is also used to explain and conceptualise the phenomenon of data analytics usage for innovation and to add a richer dimension to the secondary data obtained from literature review. The multiple case studies were conducted based on the following sub-questions:

- What is the value of data analytics to the organisation, and how is it used to derive that value?
- What is the primary focus of the organisation when starting a new data analytics project with the end goal of innovation in mind?
- How does the organisation increase the probability of achieving innovation using data analytics?
- What are the success factors for successfully innovating using data analytics?

Semi-structured open ended interviews were the primary means of data collection in the multiple case studies. In addition to this, attention was also given to the collection of secondary materials as appropriate for each case. Secondary document analysis and texts offered valuable sources of qualitative data by providing additional information as well as verifying and corroborating the information gained through interviews (e.g. a brief history of the organisation, reports, newsletters, promotional material, internet web sites and other internal publications). Moreover, extensive literature review was conducted in the area of innovation theories and innovation success factors in organisations in order to validate and compare the findings of the multiple case studies.

Finally, from the explicated success factors for using data analytics for innovation, I derived, developed and proposed a holistic model for creating data analytics enabled innovative services and products that result in significant business and customer value.

1.3 Data analytics definition

In literature, analysis of data is given different names according to the specific algorithms used, types of data, etc. For example terms such as big data analytics and advanced analytics are used to specify the specific types of data analysis performed.
In this thesis, our focus is on analysing data for creating products and services that generate significant business or customer value. Instead of focusing on specific algorithms or types of data, we believe the project case scenarios should dictate what kind of technology, algorithms, or data needs to be used. Therefore, in this thesis the term data analytics is used to refer to analysing data, regardless of big or small, for value creation regardless of the specific technologies or algorithms used.

1.4 Thesis structure

This thesis is organised into seven chapters. The remainder of the thesis will develop the argument that has been outlined in this introductory chapter.

Chapter 2 provides background information of why data analytics is important for organisations and discusses its potential for value creation and competitive advantage. Next Big data is defined and the concept of data analytics is introduced in more detail. The chapter then reviews the current state of data analytics in financial institutions in Australia by reviewing the qualitative research I conducted in collaboration with Ernst and Young. Hence, the problems that the financial institutions currently face in the use of data analytics for innovation and value generation, are discussed.

Chapter 3 starts by defining innovation and its dimensions. Next, two particular types of innovation: Disruptive innovations and big-bang disruptions - which are critical reasons for why it is crucial for financial institutions to innovate in order to stay competitive - are defined. The literature for innovation in large organisations is then reviewed and the different theoretical perspectives that are relevant in explaining the cultivation of innovation in organisations are presented. Finally, the literature on success factors for innovation in enterprises, is reviewed.

Chapter 4 defines the boundaries of the research methods that are considered for this study. It discusses the research methodology and evaluates the selection of the chosen research method adopted for this study. The chapter begins by providing a review of the philosophical assumptions of the research within the field of Information Systems, outlining the difference between positivism and interpretivism. Next, the selected research method is presented. The chapter then discusses the research design, the descriptions of the field procedures - which include survey procedures and case studies -, and the strategy utilised for data collection and analysis. Finally, the chapter accounts for the choice of grounded theory technique and describes its basic operations.
Introduction

In chapter 5 the innovative enterprises that participated in the study are presented. The account of the use of data analytics for innovation in the participating organisations are analysed and described using grounded theory techniques. The chapter proposes an enriched theoretical model that explains the success factors for use of data analytics for innovation and discusses the research findings with the existing literature. Finally, the Creative Data Analytics framework is created based on the identified key success factors. The CDA framework is proposed as a generic approach that aims to increase the probability of data analytics successfully creating products or services that create significant customer or business value.

Chapter 6 presents the research conducted to evaluate the proposed CDA framework.

Chapter 7 provides a summary of this research. The overall conclusions from the study are drawn in terms of key findings and their implications for practice and for research. The contribution of the research is presented and the chapter identifies some areas for further research. Finally, the chapter reflects on the experience of conducting this research.
2 Data – a source of competitive advantage

Introduction

The aim of this chapter is to provide an overview of data analytics and its potential for creating innovation in large enterprises that form the background for this research. First, a brief yet comprehensive definition of Big Data will be presented, followed by an overview of Data Analytics and its potential for value creation and competitive advantage. I will then discuss the qualitative research I conducted in collaboration with Ernst and Young to discover the current state of data analytics within financial institutions in Australia. The qualitative research was conducted with the aim of uncovering the challenges and problems that organisations have faced in terms of obtaining value and competitive advantage from data analytics. This chapter aims to discuss these identified problems and challenges that prevent enterprises from unlocking the true potential of data analytics.

EY has agreed to allow the IP co-created (qualitative data) in my capacity as a leader and facilitator, to be discussed in this dissertation. However, the observation and analysis of the qualitative data that is presented in this dissertation is my own work.

2.1 Big data definition

Big data is subject to immense attention from corporate leaders and academics. Normally, new technological developments first appear in technical and academic publications and the knowledge and synthesis later seeps into other avenues of knowledge mobilization, such as books. However, the fast evolution of big data technologies and the ready acceptance of the concept by public and private sectors left little time for the discourse to develop and mature in the academic domain. For immediate and wide circulation, authors and practitioners leapfrogged and published their work on big data in books and other electronic media. Therefore, one finds several books on
big data, but not enough fundamental discourse in academic publications (Gandomi and Haider; 2015).

According to Gandomi and Haider (2015) the leapfrogging of the discourse on big data to more popular outlets implies that coherent understanding of the concept and its nomenclature is yet to develop.

Big data definitions have evolved rapidly, which has created some confusion, as evident from an online survey of 154 C-suite global executives conducted by Harris Interactive on behalf of SAP in April 2012 (Gandomi and Haider, 2015). Fig. 2.1. shows how executives differed in their understanding of big data, where some definitions were focused on what big data is whereas others attempted to answer what big data does.

![Figure 2.1: Definitions of big data based on an online survey of 154 global executives in April 2012. (adapted from Gandomi and Haider (2015))](image)

Evidently, the first and most understood characteristic of big data is its size. However, other characteristics of big data have emerged recently. For example, Laney (2011) suggested that **Volume**, **Variety**, and **Velocity** (or the **Three V’s**) are the three dimensions of challenges in data management. The Three V’s have emerged as a common framework to describe big data (Chen et al., 2012; Kwon et al., 2014).

**Volume** refers to the magnitude of data. Over the past 20 years, the amount of data has increased in a large scale in various fields. According to a report from International Data Corporation (IDC), in 2011, the overall created and copied data volume in the world was 1.8ZB (≈ 1021B) (Gantz and Reinsel, 2011), which increased by
nearly nine times within five years. This figure will double at least every two years in the near future.

Information is coming from instrumented, interconnected supply chains transmitting real-time data about fluctuations in everything from market demand to the weather. In the finance industry in particular, extremely large amounts of transactional data are created everyday, containing deep information and insights into customers’ habits and needs. Additionally, strategic information has started arriving through unstructured digital channels: social media, smart phone applications and an ever-increasing stream of emerging internet-based gadgets (Lavalle et al., 2011).

Under this explosive increase in the volume of global data, the term big data is mainly used to describe enormous datasets. Big data sizes are typically reported in multiple terabytes and petabytes, where in a survey conducted by IBM in 2012, over half of the 1144 respondents considered datasets over one terabyte to be big data (Schroeck et al., 2012). In general definitions of big data volumes are relative and vary according to factors such as time and the type of data. What may be considered as big data today may not meet the threshold in the future due to the increase in storage capacities that will allow even bigger data sets to be captured. Moreover, the type of data, discussed under variety, will explain further what is meant by ‘big’. In fact, two datasets of the same size may require different data management technologies based on their type or form such as whether they are tabular data or video data. Consequently, definitions of big data also depend upon the industry, making it impractical to define a specific threshold for big data volumes.

Variety refers to the structural heterogeneity in a dataset. Advances in technology have provided the ability to utilise various types of structured, semi-structured, and unstructured data. Structured data, which constitutes only 5% of all existing data (Cukier, 2010), refers to the tabular data found in spreadsheets or relational databases. Unstructured data such as text, images, audio, and video, sometimes lack the structural organization required by machines to analyse them. Semi-structured data span a continuum between fully structured and unstructured data, and their format does not conform to strict standards.

A high level of variety, a defining characteristic of big data, is not necessarily new. Organizations have been storing (and typically not using) unstructured data from internal sources such as sensor data, and external sources such as social media for a long time. However, the emergence of new data management technologies and analytics, which enable organizations to leverage data in their business processes, is the innovative aspect.
Velocity refers to the rate at which data is generated and the speed at which it should be analyzed and acted upon. The proliferation of digital devices such as smartphones and sensors have led to an unprecedented rate of data creation and is driving a growing need for real-time analytics and evidence-based planning. Even conventional retailers are generating high-frequency data. For example, Wal-Mart processes more than one million transactions per hour (Cukier, 2010). The data emanating from mobile devices and flowing through mobile apps produce torrents of information that can be used to generate real-time, personalized offers for everyday customers. This data provides comprehensive information about customers, such as geospatial location, demographics, and past buying patterns, that can be analyzed in real time to create real customer value. These huge data feeds cannot be instantaneously handled by traditional data management systems and therefore this is where big data technologies are required. The big data technologies enable firms to create real-time intelligence from high volumes of ‘perishable’ data (Gandomi and Haider, 2015).

In addition to the three V’s, other dimensions of big data have also been mentioned. These include:

- **Veracity.** IBM coined Veracity as the fourth V, that represents the unreliability inherent in some sources of data. For instance, customer sentiments in social media are uncertain in nature since they entail human judgment, yet they contain valuable information. Therefore, the need to deal with imprecise and uncertain data is another facet of big data, that is addressed using tools and analytics developed for management and mining of uncertain data.

- **Variability (and complexity).** SAS introduced Variability and Complexity as two additional dimensions of big data. Variability refers to the variation in the data flow rates that are often not consistent and have periodic peaks and troughs. Complexity refers to generation of big data through numerous different sources. This imposes a critical challenge: the requirement to connect, match, cleanse and transform data received from different sources.

- **Value.** Oracle introduced Value as a defining attribute of big data. Based on Oracle’s definition, big data are often characterized by relatively “low value density” that means the data received in the original form usually has a low value relative to its volume. However, a high value can be obtained by analyzing large volumes of such data.
In this thesis the focus is on the value aspect of big data. While every dimension and characteristic of big data poses its own challenges and opportunities, our focus in this thesis is enabling the derivation of value from big data. While the other dimensions of big data have attracted immense attention from academia, value creation has received limited attention.

The relativity of big data volumes discussed earlier applies to all dimensions. Thus, universal benchmarks do not exist for volume, variety, and velocity that define big data and these limits depend upon the size, sector, and location of the firm and evolve over time. Also the fact that these dimensions are not independent of each other is important because as one dimension changes, the likelihood that another dimension will also change as a result, increases. However, for every organisation a ‘three-V tipping point’ exists beyond which the traditional data management and analysis technologies become inadequate for deriving timely intelligence. Therefore the Three-V tipping point is the threshold beyond which firms start dealing with big data (Gandomi and Haider, 2015).

2.2 Data analytics

Big data and analytics are inseparable for value creation as without data, there can be no analysis and without analytics, big data is merely noise. Together they offer the potential for innovation (Stubbs, 2014). According to many scholars and practitioners, Big data is likely to cause large-scale industrial and social disruption that has not been seen since the industrial revolution. Digitization has and is fuelling disruption but despite this, the fundamentals of business have not changed. Success still requires innovation, differentiation, and a relentless focus on efficient execution however, what has changed is the dynamic that information plays in this mix (Stubbs, 2014).

Information has always equated to power and entire sectors have been built on this power inequality, whether it’s at the micro-level of selling used goods through to the macro-level of financial markets. While information has always conferred advantage, the sheer volume of information available today has changed its relative contribution to success. For centuries a strong barrier to entry has been the knowledge of how the market operates and what signals to rely on. This is due to the fact that in the absence of quantitative information, experience needs to be relied on, and without experience, one is powerless (Stubbs, 2014). Big data cracks this edifice; when data becomes plentiful and accessible, the need for experience declines. This is the reason why utilising data has enabled non-banks to gain the ability to understand the market as well as or better than the incumbents in specific parts of the
finance industry. Supermarkets like the Australian brand Coles are getting banking licenses and presenting real competition to the traditional Australian banks, protected as they are by the four pillars policy (Stubbs, 2014). Similarly nonbanking institutions like PayPal are inserting themselves into the payment chain and actively dis-intermediating the banks.

Manyika et al. (2011) stated that the history of IT investment and innovation and its impact on competitiveness and productivity, strongly suggest that big data possesses similar powers that can significantly transform our lives. This suggestion is based on the fact that the same preconditions that enabled IT to power productivity are in place for big data. Based on their research Manyika et al. (2011) further claimed that there is compelling evidence that the use of big data will become a key basis of competition, underpinning new waves of productivity growth, innovation, and consumer surplus—as long as the right policies and enablers are in place. Stubbs (2014) further argues that Big data has the potential to create more efficiency, generate returns, answer “the hard questions” that no one knows the solution to, and create competitive advantage. Some of these benefits lead to internal value, such as productivity, while others lead to external value, such as revenue or may even lead to total reinvention through dynamic change (Stubbs, 2014).

These benefits have become possible because the ability to store, aggregate, combine data and then perform deep analysis has become ever more accessible as trends such as Moore’s Law in computing, its equivalent in digital storage, and cloud computing continue to lower costs and other technology barriers.

Big data is often created by dynamic sources such as intense networks of customers, clients, and companies, and thus there is an automatic flow of data that is always available for analysis. This almost voluntary generation of data can cause not only issues explained in the previous section such as data volume, velocity, variety and veracity, but also individual privacy, and ethics. Furthermore, if data points appear without anticipation or rigor of experimental design, then their incorporation in tasks like fitting a suitable statistical model or making a prediction with a required level of confidence, that may depend on certain assumptions about the data, can be challenging. On the other hand, the spontaneous nature of such real-time proactive data generation can help to capture complex and dynamic phenomena and enable data-driven decision-making provided this ability is harnessed in a cautious and robust manner.

For the enterprise market, Big Data analytics has proven its value, and examples abound. Companies such as Facebook, Amazon, and Google that in Chapter 5 of this
thesis were subjects of our multi-case studies due to their data analytics capabilities, have come to rely on Big Data analytics as part of their primary marketing schemes, product and service development, as well as a means of servicing their customers better.

For example, Amazon has leveraged its Big Data in order to create an extremely accurate representation of what products a customer should buy (a technique called recommender systems). This is accomplished by storing each customer’s searches and purchases and approximately every other piece of information available, and then applying algorithms to that information to compare one customer’s information with all of the other customers’ information.

In this way Amazon has been extracting value from large data effectively and has applied performance and depth to a massive amount of data to determine what is important and what is extraneous. The company has successfully captured the data “exhaust” that any customer or potential customer has left behind to build an innovative recommendation and marketing data element.

The results are real and measurable, and they offer a practical advantage for a customer. As an example, similar to an in-store salesperson who may recommend to a customer buying a jacket in a snowy region to purchase other items such as boots and hats; Amazon’s Big Data analytics is able to interpret trends and bring understanding to the purchasing process by simply looking at what customers are buying, where they are buying it, and what they have purchased in the past. Those datasets, combined with other public data such as census, meteorological, and even social networking data, create a unique capability that services the customer and creates significant value both for the customer and the business.

Similarly, at Facebook Big Data is utilised for critical features such as friend suggestions, targeted ads, and other member-focused offerings. Facebook is able to accumulate information by using analytics that leverage pattern recognition, data mash-ups, and several other data sources, such as a user’s preferences, history, and current activity. Those data are mined, along with the data from all of the other users, to create focused recommendations, that are reported to be significantly accurate for the majority of users.

Google also leverages Big Data, and is one of the originators of the software elements that make Big Data possible. However, Google’s approach and focus is relatively different from companies such as Facebook and Amazon. Google aims to use Big Data to its fullest extent, to judge search results, predict internet traffic usage, and service customers with Google’s own applications. From the advertising perspective,
Data – a source of competitive advantage

Web searches can be tied to products that fit into the criteria of the search by delving into a vast mine of web search information, user preferences, cookies, histories, etc.

Despite all of this potential, big data is worthless in a vacuum and its potential value is only unlocked when leveraged and analysed for driving decision making and value creation. The overall process of extracting insights from big data can be broken down into five stages (Labrinidis and Jagadish, 2012), shown in Fig 2.2. These five stages form two main sub-processes: data management and analytics. Data management involves processes and supporting technologies to acquire and store data and to prepare and retrieve such data for the purpose of analysis. On the other hand, analytics refers to techniques used to analyze and acquire intelligence from big data. Thus, big data analytics can be viewed as a sub-process in the overall process of ‘insight extraction’ from big data.

![Big Data Processes](image)

Figure 2.2: Processes for extracting insights from big data (Labrinidis and Jagadish, 2012)

### 2.3 Problems with the use of data analytics for innovation in financial sector

To understand the current state of the data analytics capabilities of the Australian large financial institutions, I in collaboration with Ernst and Young conducted a qualitative study using semi-structured open ended interviews with 12 of the largest financial institutions in Australia. These organisations are listed in the table below:

Appendix 1 shows the interview questions for this qualitative research. Below are key findings of our qualitative research:

The interviewees in our study indicated that analytics is substantially attractive to their organisation because post Global Financial Crisis, the industry has been immensely repressed and the organisations now believe analytics will provide them
with great opportunities in the form of competitive edge and more profits. Furthermore, analytics is especially attractive to these organisations because they believe it is hard to copy - as it is not in the public domain and it is not a tangible product in the hands of customers - and therefore can provide significant competitive advantage.

Therefore, to unlock the potential of data analytics, more than 90% of the organisations in our study have created a centralised analytics centre recently (between 2014-2015). Prior to the centralised centres, the analytics efforts were scattered through the organisation, which the interviewees claimed made the efforts less effective. A centralized analytics unit that is often called either a center of excellence or center of competency, makes it possible to share analytic resources efficiently and effectively. The centralized unit does not intend to replace distributed and localized capabilities; and instead the central unit is additive and is built upon existing capabilities that may have already been developed in functions, departments and lines of business (Lavalle et al., 2011). More than 85% of the banks in our study have chosen a centralized but federated model for the analytics centre due to the fact that these organisations believe a degree of understanding of the business is crucial to conducting effective data analytics. About 30% of the firms have gone a step forward and have attempted to gather together all the elements of innovation that are required to design, market, deliver and iterate on a new proposition whether it’s a radically or incrementally new product, service or process. They have brought together marketing and direct marketing, customer research, insights, and

<table>
<thead>
<tr>
<th>Organisation</th>
<th>Established</th>
<th>Market Cap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suncorp</td>
<td>1996</td>
<td>18.25 B</td>
</tr>
<tr>
<td>Allianz</td>
<td>1890</td>
<td>79.45 B</td>
</tr>
<tr>
<td>CBA</td>
<td>1911</td>
<td>136.99B</td>
</tr>
<tr>
<td>Westpac</td>
<td>1982</td>
<td>100.60B</td>
</tr>
<tr>
<td>Macquarie</td>
<td>1969</td>
<td>29.41B</td>
</tr>
<tr>
<td>IAG</td>
<td>2000</td>
<td>14.82B</td>
</tr>
<tr>
<td>QBE</td>
<td>1885</td>
<td>17.63B</td>
</tr>
<tr>
<td>HCF</td>
<td>1932</td>
<td>NA – (revenue 2.4B)</td>
</tr>
<tr>
<td>NAB</td>
<td>1981</td>
<td>78.66B</td>
</tr>
<tr>
<td>Bupa</td>
<td>1947</td>
<td>NA – (revenue 9.8B)</td>
</tr>
<tr>
<td>ANZ</td>
<td>1835</td>
<td>81.36B</td>
</tr>
</tbody>
</table>

Table 2.1: List of organisations involved in our qualitative research
demographics, product innovation, analytics and data – all of which were previously spread out across the whole group. They believe looking at innovation holistically as a centralized model is much more effective.

Furthermore, having a dedicated data analytics group may assist in fostering an analytics driven business culture. An analytics group with its own director could develop an analytics strategy and project plan, promote the use of analytics within the company, train data analysts on analytics tools and concepts and work with the IT, BI and data warehousing teams on deployment projects (Ohlhorst, 2012).

A centralized analytics unit can provide a facility for more advanced skills to come together within the organization, and provide both advanced models and enterprise governance by establishing priorities and standards using the following practices (Lavalle et al., 2011):

- Advance standard methods for identifying business problems to be solved with analytics.
- Facilitate identification of analytic business needs while driving rigor into methods for embedding insights into end-to-end processes.
- Promote enterprise-level governance on prioritization, master data sources and reuse to capture enterprise efficiencies.
- Standardize tools and analytic platforms to enable resource sharing, streamline maintenance and reduce licensing expenses.

According to the organisations in our study, the most significant problem facing their analytics efforts is the difficulty of identifying what problems/opportunities to focus their analytics endeavours on, in order to obtain business and customer value. The pain point most frequently reported was that analytics teams typically attract individuals that are proficient in the technical skills of engineering, coding, mathematics etc. but lack the ability to work effectively in a decision making team and do not think of analytics from a commercial and business value perspective. The highly technical people that are in the analytics teams most often try to find problems to solve with the new “cool” technology, as opposed to trying to find the correct technology to solve a problem that is truly valuable to the business or customers.

As Stubbs (2014) reports this is a problem of many data analytics project, where the analysts believe analytics is about data mining, visualization, machine learning,
and other functional capabilities and consequently, they are mainly interested in functionality and analytical asset creation. Model accuracy is frequently the primary benchmark for quality and once it achieves a sufficient level of quality or unveils a deeper truth, many analysts believe their work to be complete. What happens after that and how that knowledge is used to drive value is either considered irrelevant or overlooked. Typically, in organisations that have difficulty using analytics for innovation and business value, the teams responsible for analytics claim that its someone else’s responsibility to use the insights or find ways for it to be used and their role is to create insight. Insufficient attention is paid to change management and it’s taken as a given that the organization should value the insights the analysts produce (Stubbs, 2014).

Stubbs (2014) further argues that an organization can be very mature at managing technology or developing models while still being entirely incapable of innovation or value creation. An example of this is data warehousing where many organizations have mature, efficient, and highly scalable warehouses that are capable of maintaining and managing “big data.” However, many of these organizations have significantly limited knowledge in terms of how to commercialize their data assets (Stubbs, 2014). Technical skills are only one part of human capital and equally, if not more important, are behaviours and a focus on value creation. Achieving sustainable competitive advantage is a continuous endeavour and because of this, the goals will constantly shift.

Another problem is that when analytics is treated simply as a series of ad hoc activities, it is of limited value because it is focusing on activities, and not necessarily outcomes. Many analysts may be extremely proficient at using data to find answers to difficult problems and may also be experts using their technology assets however, they are less consistent about acting on the insights discovered. Consequently, the answers they find have a tendency to either disappear or be diluted because creating knowledge and answering questions is considered as success; without any attention to whether that knowledge creates any value.

Big Data offers businesses an unparalleled opportunity to extract insight into the behaviour of their customer that can in turn transform business results. However, as Bernard (2015) argues, it is much too easy to get lost in endless possibilities that can drain the businesses analytics without any definable or useful output. The danger therefore is that the analytics units and the organisations get lost in a sea of data that delivers no comprehensible value (Bernard, 2015). Similarly, Olhorst (2012) states that making a big investment to attack Big Data without first discovering
how doing so can add value to the business is one of the most serious missteps for would-be users of analytics (Ohlhorst, 2012).

All the organisations in our study believed that in order to successfully obtain business and customer value from data analytics, it is important to identify high value problems and opportunities to focus the analytics endeavours on and be specific about what problems or questions are intended to be solved using analytics. Identifying the right problems and questions that have the potential to create substantial value for business and customers - while key to analytics success - is currently considerably challenging and is not followed in most organisations. In fact, more than 80% of the organisations in our study complained about the low demand for high end analytics and stated that most of their business units are more focused on obtaining a report to validate their current agenda, than striving to solve an actual problem using analytics. There was a struggle to have the analytics units transform from having a day to day ad-hoc role in the organisation to increasingly take an innovation role.

These findings are in-line with the results of other studies in literature such as an IBM (2013) study, which found that the most effective big data and analytics strategies identify business requirements first and then leverage the existing infrastructure, data sources and analytics to support the business opportunity. In another study McAfee and Brynjolfsson (2012) concluded that companies succeed in the big data era not simply because they have more or better data, but because they have leadership teams that set clear goals, define what success looks like, and ask the right questions. They argued that Big data’s power does not erase the need for vision or human insight and on the contrary, business leaders who can spot a great opportunity, understand how a market is developing, think creatively and propose truly novel offerings, are needed.

Furthermore, Lavalle et al. (2011) suggested that companies that make data their overriding priority often lose momentum prior to the first insight being delivered, often due to the fact that a data-first approach can be perceived as taking too long before generating a financial return. By narrowing the scope of these tasks to the specific subject areas needed to answer key questions, value can be realized more quickly, while the insights are still relevant. Lavalle et al. (2011) argue that organisations should implement analytics by first defining the insights and questions needed to meet the critical business objective and then identifying those pieces of data needed for answers.
Bernard (2015) similarly argues that to use Big Data to improve the business and gain a competitive advantage, there is a need to understand what the big strategic objectives are. Once these objectives are defined, then it is critical to isolate the critical and high-value questions that need to be answered, in order to identify the data – small or big – that will provide the answers to those questions. Starting with strategy, therefore, allows the organisation to identify the strategic information needs.

Bernard (2015) believes that in order to cut through the chaos, confusion and sheer volume of data that can or does exist, it is important to ‘start with strategy’. Instead of starting with the data, starting with the business objectives and the aim that is to be achieved, will direct the efforts toward questions that need to be answered. This will immediately narrow data requirements into manageable areas (Bernard, 2015).

In conclusion, our study shows that the technology is not the most important challenge that organisations face when adopting analytics. The adoption barriers that organizations face most are managerial and cultural rather than related to data and technology. The leading obstacle to widespread analytics adoption is lack of understanding and difficulty of knowing how to use analytics to improve the business and the challenge of having the creativity needed to derive innovative services and products using data analytics that create value. A survey of nearly 3000 executives, managers and analysts working across more than 30 industries and 100 countries, conducted by MIT Sloan Management Review and IBM Institute also reported this same finding (Lavalle et al. 2011).

2.3.1 Current data analytics model

Cross Industry Process for Data Mining (CRISP DM) is one of the most used reference process models for data analytics. All the financial institutions in our study use a version of CRISP DM model for their data analytics projects. CRISP DM was created in 1996 on the basis of best industry practices. According to CRISP DM, the life cycle of a data analytics project consists of six phases, shown in Figure 2.3. In the life cycle, the sequence of the phases is not rigid and moving back and forth between different phases is always required. The outcome of each phase determines which phase, or particular task of a phase, has to be performed next. The arrows indicate the most important and frequent dependencies between phases. Below I will review each of the stages in this model:
Business understanding

This initial phase focuses on understanding the project objectives and requirements from a business perspective, then converting this knowledge into a data analytics problem definition and a preliminary plan designed to achieve the objectives. As evident, the CRISP DM model provides no framework or methodology for identifying these project objectives and requirements, and assumes that they are already devised. We have already discussed that identifying the objectives of the data analytics projects in order to ensure business and customer value can be obtained, is the most difficult challenge of data analytics projects. Consequently, there is an obvious gap in this current model for data analytics.

Data understanding

The data understanding phase begins with primary data collection and will subsequently involve activities that enable the familiarisation of the analyst with the
data, leading to identifying data quality problems, discovering first insights into the data, and/or detecting interesting subsets to form hypotheses regarding hidden information.

**Data preparation**

The data preparation phase involves all activities that are required to construct the final dataset [data that will be utilised in the modeling tool(s)] from the initial raw data. Data preparation tasks are commonly performed several times and not in any prescribed order. These tasks comprise of activities such as table, record, and attribute selection, as well as transformation and cleaning of data for modeling tools.

**Modeling**

This phase involves applying and selecting various modeling techniques, and calibrating their parameters to optimal values. Commonly, there are multiple techniques for the same data analytics problem type and some techniques have specific requirements on the form of data. Consequently, iterating back to the data preparation phase is often required.

**Evaluation**

This stage of the data analytics project involves constructing a high quality model (or models) from a data analysis perspective. Before proceeding to final deployment of the model, it is critical to evaluate the model(s) entirely and review the steps executed to create it, in order to ensure that the model achieves the business objectives adequately. A fundamental objective of this phase is to determine if there are any critical business issues that have not been considered adequately. At the end of this phase, a decision on the use of the data analytics results should be reached.

**Deployment**

Typically the generation of the model is not the aim of the project, and even if the purpose of the model is to increase knowledge using data, the knowledge gained is required to be organized and presented in a way that the customer can use it. It frequently involves applying “live” models within an organization’s decision making processes such as real-time personalization of web pages or continuous scoring of marketing databases. Depending on the requirements, the deployment phase can be as simple as generating a report or as complex as implementing a repeatable data analytics process across the enterprise.
The CRISP DM model of data analytics assumes that the business problem is already identified and provides no means or approaches to identifying a high value business problem that can result in innovative products and services. Therefore, this model does not assist in solving the problem that the financial institutions currently have with enabling innovation through the use of data analytics.

**Conclusion**

For financial institutions, with no physical products, data – the source of information - is the most valuable asset they possess and one that can be the enabler of unprecedented competitive advantage and innovation. Data contains valuable insights into customers’ preferences, allowing it to be used as a basis for the creation of products and services that remedy emerging problems or opportunities in a timely and more competitive manner.

The use of analytics is thus becoming a key foundation for competition and growth for individual firms, and it will underpin new waves of productivity, growth, and consumer surplus. This is because it is not the data itself or the amount of data that creates the value, but our ability to analyse vast and complex data sets beyond anything we could ever do before. (Bernard, 2015).

While obtaining unprecedented value and competitive advantage are expected from data analytics, our qualitative research aligned with the results of other studies in the literature revealed that the finance sector is struggling to obtain this value. According to our study this lack of ability to use the potential of data analytics is mainly due to the fact that organisations are unable to determine what projects to focus on in order to create innovative products and services that create significant business or customer value.
3 Innovation

Introduction

This thesis aims to cultivate innovation in enterprises using data analytics. Innovation is therefore a critical factor in this research and consequently in this chapter I review the literature of innovation in enterprises. The innovation literature is highly fragmented drawing from multiple disciplines and is studied using a wide range of research methodologies. There are also multiple levels of analysis and dimensions that further add to the complexity. The aim of this chapter is to integrate this disparate literature of innovation and provide an overview of first the dimensions of innovation and second the success factors of innovation in enterprises.

This review of existing literature and theories allows me to intensify my theoretical sensitivity, by providing me with insights and the ability to give meaning to data, as well as the capacity to understand and separate critical information from a mass volume of data. It is this theoretical sensitivity that will allow me to develop a theory that is grounded, conceptually dense and well integrated. Furthermore, the existing innovation literature review will act as a point of comparison of the findings of this study, allowing to determine which findings are in-line or in contrast with pre-conducted studies.

3.1 Innovation definition

Innovation is widely believed to be a critical source of competitive advantage in an increasingly changing environment (Dess and Picken, 2000; Tushman and O’Reilly, 1996). Both researchers and practitioners realize the importance of innovation as witnessed by thousands of academic papers and numerous business rankings and indices. According to management scholars, innovation capability is the most important determinant of firm performance (Mone et al., 1998). This is because although a firm with a higher business growth performance is considered to have a competitive advantage due to its valuable, unique, and difficult to imitate resources
Innovation and capabilities, the sustainability of its competitive advantage might depend on its innovative capacity (Porter, 1980; Barney, 1991).

There are various definitions of innovation, each emphasizing a different aspect of the term innovation. The first definition of innovation was coined by Schumpeter in the late 1920s (Hansen and Wakonen, 1997), who stressed the novelty aspect. In his 1911 book “The Theory of Economic Development”, Joseph Schumpeter renounced traditional economics and stated that the norm of healthy economics and fundamental reality of economic theory and practice did not lie in balance and optimisation but rather in dynamic imbalance caused by innovative activity. He argued that profit may be derived not only from alteration of price and avoidance of cost but also from radical placement of output goods.

According to Schumpeter, innovation is reflected in novel outputs: a new good or a new quality of a good; a new method of production; a new market; a new source of supply; or a new organizational structure, that can be summarized as ‘doing things differently’. However, as Hansen and Wakonen state, “it is practically impossible to do things identically” (Hansen and Wakonen, 1997, p. 350), which by definition makes any change an innovation. Although Schumpeter clearly positioned his definition of innovation within the domain of the firm and outlined its extent as product, process, and business model, there are continuing debates over various aspects of innovation: its necessity and sufficiency (Pittaway et al., 2004), its intentionality (Lansisalmi et al., 2006), its beneficial nature (Camison-Zornoza et al., 2004), its successful implementation (Hobday, 2005; Klein and Knight, 2005), and its diffusion (Holland, 1997) to qualify as innovation.

Many scholars regard innovation as an integral activity that involves the whole firm and conditions its behaviour (Martinez-Roman et al., 2011; Yam et al., 2011) and therefore relate an innovation to the ability of the firm to seek new and better ways to identify, acquire, and implement ideas and tasks in an organisation (North and Smallbone 2000; Calantone et al., 2002; Blumentritt and Danis, 2006; Brem and Voigt, 2009; Hjalager, 2010). Consequently, in order for innovation to exist in a firm, it is necessary to cultivate both the external and internal environments and success factors, and the driving forces. This is to understand the innovation potential and the innovation activities required (Hjalager, 2010; Martinez-Roman et al., 2011).

In 2010, Crossan and Apaydin composed a comprehensive definition of innovation according to which innovation is: “production or adoption, assimilation, and exploitation of a value-added novelty in economic and social spheres; renewal and enlargement of products, services, and markets; development of new methods of
production; and establishment of new management systems. It is both a process and an outcome”. This definition captures several important aspects of innovation: it includes both internally conceived and externally adopted innovation (‘production or adoption’); by including application (‘exploitation’), it highlights innovation as more than a creative process; it emphasizes intended benefits (‘value-added’) at one or more levels of analysis; it leaves open the possibility that innovation may refer to relative, as opposed to the absolute, novelty of an innovation (an innovation may be common practice in other organizations but it would still be considered as such if it is new to the unit under research); and it draws attention to the two roles of innovation (a process and an outcome). Due to the comprehensibility of Crossan and Apaydin’s definition we use this innovation definition in this thesis.

3.2 Innovation overview

It was important to establish the definition of innovation that we are using in this thesis as innovation is a broad term with multiple meanings. Innovation literature draws on theories from a variety of disciplines and has been studied using a wide range of research methodologies. Multiple levels of analysis and dimensions, and inconsistent operationalization of the primary constructs, which in turn has led to mixed empirical results, have further complicated this synthesis.

Most literature on innovation, especially in highly cited papers, are surprisingly purely descriptive. In the cases when a theory was utilised, the theories most frequently used were learning and knowledge management theories, followed by network theories, and economic theories. Institutional theory, resource-based view (RBV), and adaptation theories were also used but less commonly. Please refer to Table 3.1 (Crossan and Apaydin, 2010) for key references.

Level of analysis (multi-level, macro (economy/industry/market), organisation, micro (group/team/individual)) also affects the distribution of theories. Network, learning, and knowledge theories are used across all levels while economic theories are mostly used at the economy or societal level, and evolutionary economics is used evenly across macro levels. Moreover, resource-based view and adaptation theories are used at the organizational level, and psychological theories as expected are applied at the individual level. Table below presents the theories used in highly sited papers and organises them by levels of analysis.
<table>
<thead>
<tr>
<th>Innovation</th>
<th>Multilevel</th>
<th>Macro (economy/industry/market)</th>
<th>Organization</th>
<th>Micro (group/team/individual)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Haunschild and Miner (1997)</td>
<td>(contingency)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pil and Macduffie (1996)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Porter (1998)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Westphal et al. (1997)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource-based view and dynamic capabilities</td>
<td></td>
<td></td>
<td>Christmann (2000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lei et al. (1996)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Teece (1998)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tidd et al. (1997)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Eisenhardt and Tabrizi (1995)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Grindley and Teece (1997)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lam (2005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>McGrath (2001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Powell (1998)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Powell et al. (1996)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tushman and O’Reilly (1996)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sorensen and Stuart (2000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(interactionist)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Theories used in highly cited papers (Crossan and Apaydin, 2010)
In this thesis we focus on organisational innovation. Our focus on organisational innovation (firm, group, and individual levels of analysis) was driven by the purpose of being practical in our orientation by concentrating on elements that are debatably within the control of the firm. Despite the fact higher level models would be more comprehensive, they would inevitably include industry, national, or global levels, which are arguably beyond the individual firm’s control. Therefore, by targeting at the firm level, we can provide a practical basis on which managers can build structures and systems that would enable innovation (specially through the use of data analytics) within a firm. It was also important to isolate the leaders’ influence from organizational level factors because the mechanisms for the leadership’s connection with the rest of the innovation process have not been explicit, even though leadership for innovation has been a subject of research. In addition, in line with the innovation definition we have adopted, we intend to delineate the difference between innovation processes and outcomes: the former clearly precedes the latter and should be separated to avoid circular arguments.

Since the aim in this thesis is to identify actionable success factors that are within the boundaries of organizational and individual power, we limit our level of analysis to the firm level (the two right columns of the table). Organisational innovation can be divided into two sequential components: 1. dimensions of innovation (which itself has two categories of innovation as a process and innovation as an outcome) and 2. success factors of innovation that look at innovation leadership. The role of leadership at all levels of an organization, although sometimes implicit, is critical for innovation as a process and maintaining its momentum until innovation as an outcome ensues. Adoption of this sequential view helps bring to light the often missed causal interconnectedness between these three components of innovation.

### 3.3 Dimensions of innovation

In this section we discuss the first sub-component of organisational innovation, which is the dimensions of innovation. The definition adopted in this thesis provides the first relationship in the dimensions of innovation, which is innovation as a process will always precede innovation as an outcome. Thus, these two roles of innovation are separated. A total of ten dimensions of innovation surfaced from the literature and by revising the various dimensions, it was deduced that these dimensions could be organized into two categories: those relating to innovation as a process and those concerning innovation as an outcome. The former answered the question ‘how’ while the latter answered the question ‘what’.
3.3.1 Innovation as a process.

Driver and source

The driver and source dimensions can be either internal or external. An internal driver of the innovation process can be available knowledge and resources, while an external driver would be a market opportunity or imposed regulations. In terms of the source dimension, an example of an internal source of innovation is ideation, while an external source of innovation is adoption of innovation invented elsewhere.

External sources of innovation are exploited by many incumbents, where they use Mergers and Acquisitions (M&A) as an innovation enabler and disruption mitigator. In the context of innovation, M&A can be conceived of as mechanisms to accomplish several distinct roles including but not limited to (a) filling out a product line; (b) reducing time to market; (c) obtaining scale economies in research and bringing down the unit cost of R&D; (d) enhancing the rate of innovative output of an organization; (e) enabling absorption of new technology to enhance manufacturing or operations capability; (f) helping organizations enhance an existing capability (e.g. in an area in which the organization is already active); and (g) helping organizations build a new innovation capability (e.g. in an area in which the organization is currently not active or marginally active) (Dodgson et al., 2014).

A significant number of studies have focused on investigating whether and how, engaging in an acquisition may lead to higher innovation. This issue has attracted considerable research interest because, although many acquirers’ ultimate goal is innovation and some studies do report a positive impact of engaging in acquisitions on the innovation outcome of firms (Capron and Mitchell, 1998; Desyllas and Hughes, 2010), other studies show the opposite effect. As an example, several studies report a negative relationship between acquisition intensity and the rate of internal innovation, because acquisitions require time and attention for the extensive preparation, negotiations, and integration activities involved in the process (Hitt et al., 1990; Hitt et al., 1996). Another possible cause of this negative impact on innovation is managers overestimating their ability to manage an acquired business (Hitt et al., 1991). Chaudhuri and Tabrizi (1999) stress many common difficulties that lead acquisitions to have disappointing results for acquiring firms. At the individual level, research has revealed that acquisitions may have a negative impact on the productivity of inventors (Kapoor and Lim, 2007) and that frequently key inventors leave the acquired firm (Ernst and Vitt, 2000). Due to these conflicting results, a substantial body of research has focused on investigating the conditions under which acquisitions lead to an improvement on innovation. The factors identified by literature in
Innovation

this area can be classified into three main categories of dyadic-level factors (such as the relatedness of the acquired knowledge-base), firm-level factors (motivation, characteristics of underlying knowledge, firm characteristics), and factors related to acquisition implementation (Dodgson et al., 2014).

Often the incumbents wait to see if an innovation implemented by a startup or company is successful in the market and if so they will then try to acquire the company. However, there are problems with this approach such as the innovator may not be willing to be involved in a M&A (such as in the case of Facebook) and also as discussed before the probability of success of M&A is reasonably low. Therefore, despite the fact that M&A remain an important means of creating innovation capability in firms, having an internal innovation capability is crucial. In this thesis the focus is on the firm’s internal innovation capabilities, that can be enhanced by the exploitation of data analytics.

Locus, view and level

The locus dimension defines the extent of an innovation process: firm only (closed process) or network (open process). The view dimension considers how the innovation process starts and develops; whether it is top-down or bottom-up. And finally, the level dimension delineates the split between individual, group, and firm processes.

3.3.2 Innovation as an outcome.

The distinction between innovation as a process and as an outcome is sometimes blurred and as Sood and Tellis (2005) indicate, lack of clarity in separation of these two facets of innovation may be fundamentally problematic. This problem is especially evident when innovation outcomes are confused with market performance (e.g. new entrants that displace incumbents with disruptive technologies) such that researchers risk asserting premises that are true by definition. Dimensions concerning innovation as an outcome should answer the questions ‘what’ or ‘what kind’ and these dimensions include referent, form, magnitude, type, and nature.

In terms of the type dimension of innovation, Gopalakrishnan and Damanpour (1997) separate technical and administrative innovations that reflect a more general distinction between social structure and technology. Technical innovations comprise of products, processes, and technologies utilised to produce products or services that are directly related to the basic work activity of an organization. On the other hand, administrative innovations are not directly related to the basic work activity
Innovation

and instead are related to its managerial aspects such as organizational structure, administrative processes, and human resources.

In regards to the dimension of form, scholars have suggested three distinctions, which are: product or service innovation, process innovation, and business model innovation. Product/service innovation is “the novelty and meaningfulness of new products introduced to the market in a timely fashion” (Wang and Ahmed, 2004, p. 304). Novelty can also be a dependant of the referent dimension where a product or service can be new to the company (Davila et al., 2006), the customer (Wang and Ahmed, 2004), or the market (Lee and Tsai, 2005). Process innovation is the “introduction of new production methods, new management approaches, and new technology that can be used to improve production and management processes” (Wang and Ahmed, 2004, p. 305) and therefore process innovation is an internal phenomenon that technically results in the referent being the firm itself. Business model innovation is “how a company creates, sells, and delivers value to its customers” (Davila et al., 2006, p.32), regardless of whether it is new to the firm, customer, or industry. Damanpour and Aravind (2006) conducted an extensive meta-analysis of the empirical studies of the effects of organizational and environmental success factors on product and process innovation, that revealed most success factors do not differentiate between these two types of innovation outcomes. This suggests that product and process innovations are complementary and not distinct from one another.

The referent dimension defines the newness of innovation as an outcome; where the innovation can be new to the firm, to the market it serves, or to the industry. The referent dimension is associated to several of the other dimensions. For instance, the referent and magnitude dimensions are visibly interconnected: while incremental innovation such as continuous improvement initiatives may be new to the firm, more radical innovation will be associated with the market and even industry.

The magnitude dimension specifies the degree of newness of the innovation outcome with respect to an appropriate referent and is typically divided into incremental and radical innovations (Gopalakrishnan and Damanpour, 1997). The latter is sometimes termed as ‘revolutionary’, ‘disruptive’, ‘discontinuous’, or ‘break-through’ (Freeman, 1974; Garcia and Calantone, 2002; Tushman and Anderson, 1986). Radical innovation stimulates fundamental changes and a clear departure from existing practices in the organization, while incremental innovation represents a variation in existing routines and practices (Damanpour, 1991; Dewar and Dutton, 1986; Ettlie et al., 1984). The absolute and relative magnitude of innovation outcomes have been a subject of debate.
Two special kinds of radical innovations are disruptive innovations and big-bang disruptions. These two special kinds of innovation have a substantial disruptive effect on all industries including the finance industry and pose dramatic threats that make the necessity to innovate even more crucial. Due to the importance of these two disruptions, I will discuss them in further details in the next section.

### 3.3.2.1 Disruptive innovation

**Definition**

Disruptive Innovation Theory, advanced by Christensen (1997, 2006) Christensen and Bower (1996) and Christensen and Raynor (2003), was created based on a series of prior technological innovation studies. In 1997, Christensen published his influential book entitled The Innovator’s Dilemma, which made him renowned in the study of technological innovation in commercial enterprises. The book, which became a best seller at the time, articulated the basic theory of disruptive technology that describes a process by which a product or a service initially conceives in rudimentary applications at the bottom of a market and then relentlessly moves up market, and ultimately displaces the established competitors.

The theory takes the view that companies tend to innovate faster than their customers’ needs evolve, and due to this, eventually most organisations have established products or services that are too sophisticated, expensive, and complicated for many customers in their market. This, as Christensen claims, is the result of companies pursuing “sustaining innovations” at the higher tiers of their markets because charging the highest prices to their most demanding and sophisticated customers at the top of the market will let them achieve the greatest profitability and has historically helped them succeed.

Following this approach however, inadvertently opens the door to “disruptive innovations” at the bottom of the market. A disruptive innovation will enable a large number of consumers at the bottom of a market, to access a product or service that was previously only accessible to consumers with excessive funds or skills (Christensen and Raynor, 2003).

**Characteristics**

In disruptive innovations, the innovations that harm established companies are often not radically new or difficult from a technological point of view. They do however have two critical characteristics:
1. They usually present a different package of performance attributes that are at least at the inception not valued by existing customers.

2. The performance attributes that the existing customers do value, improve at such a rapid rate that the new technology can later invade those established markets.

Only at the second stage mentioned above, will mainstream customers desire the technology, at which point, for established suppliers it is frequently too late to enter the market, as the pioneers of the new technology dominate the market.

Examples

According to Christensen, failure of leading companies to stay at the top of their industries when technologies or markets change, is one of the most consistent patterns in business and examples of this include:

- Goodyear and Firestone entered the radical-tire market extremely late.
- Xerox let Canon create the small-copier market.
- Sears enabled the entrance of Wal-Mart.
- IBM dominated the mainframe market but missed the emergence of minicomputers (that were technologically much simpler than mainframes) by many years.
- Digital Equipment dominated the minicomputer market with innovations like its VAX architecture but missed the personal-computer market almost entirely.
- Apple computer pioneered the world of personal computing and established the standard for user friendly computing but brought its portable computer with a five year lag behind the leaders.

How to spot it

To identify the distinctions in the impact of certain kinds of innovations on a given industry, Christensen (1995) introduced the concept of performance trajectories that is the rate at which the performance of a product or service has improved and is expected to improve over time. This methodology works by plotting the product or service performance in a simple graph, with the vertical axis showing the performance as defined in mainstream markets and the horizontal axis showing time.
Innovation (figure 3.1.). The first step is to draw a line that depicts the level of performance and the trajectory of performance improvement that customers have historically desired and are likely to expect in the future. Next the estimated initial performance of the new innovation will be located and if the innovation is disruptive, the point will lie far below the performance demanded by current customers.

If knowledgeable technologists believe that the probable slope of performance improvements of the disruptive innovation or technology might progress faster than the market’s demand for performance improvement, then that innovation, which does not meet customers’ needs today, has the ability to address them in the future. And therefore, the innovation is strategically critical.

3.3.2.2 Big-bang disruption

Disruptive innovation assumes that disrupters start with a lower-priced, inferior alternative that is initially only attractive to the least profitable customers, and therefore providing the incumbent business with the time needed to start reacting and developing its own next-generation products or/and services. Big-bang disruption however, does not have any such restriction; and it can happen in what may appear as an instant.

Definition
Today, entire product lines and whole markets are being created and destroyed in extremely short amounts of time, that make the impact of these kinds of innovations unprecedented. Such disruption is difficult to fight, once launched. Downes and Nunes (2013), called this type of disruption “big-bang disrupters” and described the phenomenon as the kind of disruption that does not follow conventional strategic paths or typical patterns of market adoption and is different from more traditional innovations not just in degree but in kind. These innovations are not only cheaper than established offerings, but also they are more inventive and better integrated with other products and services – with many of them exploiting consumers’ growing access to product information and the ability to contribute and share it.

The extremely fast adoption speed of these technologies is a function of the considerably accurate market information that are enabled by many new trends such as the ease of access of customers to mobile devices and free databases of user-generated reviews – to find the best price and quality and the next new, innovative product or service.

The most significant challenge that big-bang innovations pose to the incumbents including those in the finance sector, is that they typically combine existing technologies that do not even seem related to the incumbent’s offering to achieve a dramatically better value proposition. Other characteristics of big bang disrupters is that they do not typically share the incumbents’ approach to solving customer needs and may not even see the incumbents as competition. In addition, they frequently are not trying to directly compete with the incumbents’ product line, and therefore are not attempting to find a way to offer slightly better price or performance with hopes of gaining short-term advantage. Conversely, they often create a valuable proposition for customers hoping to attract them to do a business that is entirely different from the incumbents. Competitors like this, blindside incumbents as they do not simply create the need for faster strategy formulation and execution, and more effective operations and instead they create the need for entirely new innovation, strategy and go-to-market approaches.

**Characteristics**

Downes and Nunes (2013), described big-bang disrupters as having three main devastating features:

1. **Unencumbered development**: big-bang disruptions are often the creation of rapid-fire, low cost experiments on fast-maturing, ubiquitous technology platforms and they do not require budget approval or vetting before the start of
their development. Disruptive technologies are becoming increasingly cheaper to manufacture and deploy and therefore innovators can experiment with new applications while imposing minute risks to investors and abandoning prototypes that do not quickly prove popular.

2. Unconstrained growth: Big-bang disruptions do not follow the regular product life cycle that is described by Everett Rogers’s classic bell curve of five distinct customer segments – innovators, early adopters, early majority, late majority, and laggards. Big-bang disruptions only have two segments: trial user, who typically contribute to the product development, and then all remaining consumers. Therefore the new product cycle can be simplified into three basic stages of development, deployment, and replacement.

The adoption of disruptive innovation is no longer defined by crossing a marketing chasm and instead innovators repeatedly get it wrong, before creating a product that dominates the entire market.

3. Undisciplined strategy: Big-bang disrupters do not follow the typically known competitive strategy. In their widely popular 1995 book “The Discipline of Market Leaders”, Michael Treacy and Fred Wiersema suggest that businesses should align strategic goals along only one of three value disciplines: cost (“operational excellence”), constant innovation (“product leadership”), or customised offerings (“customer intimacy”). They further suggested that failing to choose only one of these areas will result in a disjointed and disordered business. Similarly, Michael Porter offered his three generic strategies for achieving competitive advantage and warned against pursuing more than one.

Big-bang disrupters however, are thoroughly undisciplined and start with having a better performance at a lower price and greater customization. From inception, they compete with mainstream products on all three value disciplines.

The main reason that big-bang disruptions have a better performance and still are able to maintain a lower price, is the faster, cheaper and smaller computing power (predicted by Moore’s law) that is now deployable on global scale and delivered through the cloud to inexpensive mobile devices. The major cost drivers of products and services are: the parts and manufacturing, the embedded technologies and intellectual property, and a prorated share of development costs. By continually and significantly lowering all of these cost drivers simultaneously, today’s technologies make it feasible to sell new products and services at a lower price than the inferior alternatives that they displace. Due to this Dennings (2014) stated: “if Christensen’s disruption from below was scary, big bang disruption can be downright terrifying.”
Example

An example of big bang disruptions, is the catastrophic disruption of navigation-product makers like TomTom, Garmin and Magellan by free navigation apps that are now preloaded on every smartphone (Downes and Nunes, 2013). These products are not only cheaper but are better than the stand-alone devices those companies sell and are also constantly improving with new versions distributed automatically through the cloud, owing to the robust platform provided by the iOS and Android operating systems. Here not only the competition did not come from companies in the same industry or with the same business model, but also the new technology did not enter at the bottom of a mature market and then follow a carefully planned march through larger customer segments. Users made the switch in a matter of weeks and this did not just include the least profitable or “underserved” customers, consumers in every segment defected simultaneously and in large groups (Downes and Nunes, 2013).

How to spot it

In their book “Big Bang Disruption”, Downes and Nunes (2014) offered a helpful conceptual framework for thinking about big bang disruption based on four phases: a) singularity; b) big bang; c) big crunch; and d) entropy that are represented in figure 3.2. The features of the graph are the length of the period of singularity (i.e. experimentation), the shortness of the exploitation period (the big bang itself) and the rapid decline or big crunch, that can be almost as steep as the big bang.
Innovation

Some industries such as the German small and medium scale enterprises have not been greatly disrupted by big bang disruption, in part because they are in highly specialised industry segments that are not easy to disrupt (Dennings, 2014). Other sectors or geographic regions such as the finance sector have escaped as a result of being protected by government regulation. However, clearly business models that can be transformed with digital technology – such as in the case of the financial industry that has no physical products - are the most vulnerable. Even though the financial institutions appear to have their business models and sector currently protected by regulations, they may be receiving only a temporary reprieve. The protection may be trapping these incumbents into a false sense of security and unpreparedness for the moment when an unexpected breakthrough in price, performance and customization occurs or when public pressure breaks down the protective regulatory barriers. Again, this was seen in the case of Uber disrupting the highly regulated taxi industry, and the public pressure forced governments to bring down the barriers to entry and acknowledge Uber as a legal entity. The two mentioned disruptions especially the big bang disruptions make it necessary for the financial institutions to innovate in order to survive the waves.

3.4 Innovation success factors

In this section we will review the second sub-section of organisational innovation literature, which is the innovation success factors. The success factors of innovation discussed in existing literature can be divided into four distinct meta-theoretical constructs: innovation leadership, managerial levers, business processes, and customers. While the customers category could have been placed as a subcategory of managerial levers, due to its importance and critical role in innovation we devoted a separate section to it. Each of the three constructs of innovation leadership, managerial levers, and business processes can be supported by a distinct theory: the upper echelon theory, the dynamic capabilities theory, and process theory respectively. Customers construct can also be supported by the dynamic capabilities theory.

Traditionally the upper echelon theory (Hambrick and Mason, 1984) has been utilised to connect agents’ characteristics and behaviours with organizational outcomes; however, it is unable to sufficiently cover managerial levers and business processes. Conversely, dynamic capabilities research (Eisenhardt and Martin, 2000; Prahalad and Hamel, 1990; Teece et al., 1997) is concerned with organizational resources and capabilities but it is not able to entirely incorporate the role of the agent or to investigate how organizational processes transform inputs into outputs, which is the
Innovation realm of organizational process theory (Engestrom, 1993; Van de Ven and Poole, 1995). Consequently, each meta-construct requires a distinctly separate theoretical basis.

3.4.1 Innovation leadership

Innovation Leadership is a meta-construct that associates individual and group level variables. According to the findings of various studies, executives describe about 5–20 per cent of variance in company profitability (Crossland and Hambrick, 2007). Their influence on innovation was captured in a special issue of The Leadership Quarterly (2004, Vol. 15, No. 1) that was dedicated to the subject of leadership for innovation. Mumford and Licuanan (2004) provided a summary of the findings of this issue by confirming the multiple roles of leaders. The leaders’ support and guidance is vital in promoting innovative efforts at the initial creative stage, because it contributes to effective interactions among group members (West et al., 2003). Additionally, they are able to create conditions for the subsequent implementation of innovation (Mumford and Licuanan, 2004).

Upper echelon theory suggests that leaders’ behaviours are a function of their values, experiences, and personalities (Hambrick and Mason, 1984). Further to this, Mumford et al. (2002) argue that to lead creative efforts, leaders must possess significant technical and professional expertise and creative skills, in addition to the ability to process complex information. Furthermore, they must have the motivation to exercise this ability. This motivation, according to Sternberg et al. (2003), partly depends on leaders’ perception of environmental threats and opportunities. The leaders’ ability and motivation to innovate can be consolidated into two groups of factors: individual (CEO) and group (Top Management Team and Board Governance).

On the individual level, these factors include tolerance of ambiguity (Barron and Harrington, 1981; Patterson, 1999), self-confidence (Barron and Harrington, 1981), openness to experience (George and Zhou, 2001; Patterson, 1999; West, 1987), unconventionality (Frese et al., 1999; West and Wallace, 1991), originality (Patterson, 1999; West and Wallace, 1991), rule governess (Frese et al., 1999; Simonton, 1991), authoritarianism (Simonton, 1991), independence (Patterson, 1999; West, 1987), proactivity (Seibert et al., 2001), intrinsic (versus extrinsic) attribution bias (Frese et al., 1999; West, 1987), determination to succeed (Amabile, 1983), personal initiative (Frese and Zapf, 1994), and managerial tolerance of change (Damanpour, 1991).
For disruptive innovations specifically, since most senior managers were trained in conventional business programs to manage organizations that served established markets with well-defined product lines, they may not understand the promise of disruptive innovation as their views of the world are deeply entrenched and largely shaped by their current experiences (Henderson, 2006). Therefore, an additional team at the corporate level is required to be predominantly responsible for collecting disruptive innovation ideas and putting them into implementation (Christensen and Raynor, 2003). Moreover, long-term-oriented, subjective-based incentive plans should be adopted instead of short-term-oriented, formula-based incentive plans for key executives (Govindarajan and Kopalle, 2006), so that the senior managers will not be restricted by inflexible incentives and avoid the risks of disruptive innovation.

In addition to individual level, at group level there are also identified variables for innovation success. Upper echelon theory proposes that composition and characteristics of the top management team (TMT) generates a stronger explanation of organizational outcomes than a leader’s characteristics alone, including amount of education and age (Bantel and Jackson, 1989; Hambrick and Mason, 1984), tenure (Bantel and Jackson, 1989; Finkelstein and Hambrick, 1990), diversity of background and experience (Bantel and Jackson, 1989), and extra-industry ties (Geletkanczy and Hambrick, 1997). Further studies have studied board diversity in terms of occupational background (Goodstein et al., 1994), institutional shareholding (Kochhar and David, 1996), and executive stock option (Sanders and Hambrick, 2005). The Innovation Leadership construct is connected to organizational and contextual factors through Managerial Levers that play direct and indirect roles in enabling innovation. Leaders implement deductive innovation strategies (Regnér, 2003) through direct measures such as decisions and actions taken by leaders to deliver innovation. Senior executives exercise indirect leadership (Jansen et al., 2009) to guide innovation champions at the middle management level in their implementation of Business Processes that support innovation.

The role of middle managers is particularly important for disruptive innovations, because most strategic proposals take their fundamental shape at the lower levels of hierarchical organizations. As middle managers typically have the most to lose in any basic change, they are likely to allocate their resources to sustaining innovations that strengthen their current fiefdom and careers (Christensen and Raynor, 2003; Denning, 2005). Also, there may be different performances between founders and professional managers in disruptive innovations since founders have an advantage in tackling disruption because they not only exert the necessary political clout but also have the self-confidence to override established processes (Christensen and Raynor, 2003). The manager aspect discusses the key roles of both senior and middle
managers as well as the different performances between founders and professional managers, explaining why incompetence of managers could be a main obstacle to a promising disruptive innovation.

In the next section we look at Managerial Levers that link individual or group determinants with organizational factors and provide the necessary connection between leadership intentions and organizational results.

3.4.2 Managerial levers

Managerial Levers is a meta-construct that establishes the firm-level variables supporting innovation. I commence with a discussion of the dynamic capabilities theory, which supports the managerial levers construct, and hence describe the five sets of managerial levers. The construct of managerial levers can be most appropriately conceptualized using the theory of dynamic capabilities (Eisenhardt and Martin, 2000; Prahalad and Hamel, 1990; Teece et al., 1997), which is a dynamic strain of the resource-based view (Barney, 2001) that utilises the evolutionary economics (Nelson and Winter, 1982; 2002), according to which different resource bases among firms provide the source of “variation” for innovations. The marketplace will subsequently “select” the new products. The task of the organisation is to transfuse the exploitation of the existing resources while searching for new opportunities (exploitation). Constant changes in the environment and competitive landscape however, may lead to “creative destruction” (Schumpeter, 1934) of the currently valuable resources. Consequently, a firm, in addition to exploiting existing resources, is required to develop new and valuable resources and capabilities (Rumelt, 1984), that take time, investment, and managerial effort (Dierickx and Cool, 1989).

Hence, the concept of dynamic capabilities (i.e. innovation capabilities) is used to explain the ability of a firm to develop its resources and competencies to adapt to changing business and market environments that are described as being strategically responsive (Teece et al., 1997).

Depending on the strategic orientations and the market conditions of the firm, the term innovative capability is looked at from different levels and from different perspectives, (Guan and Ma, 2003; Martinez-Roman et al., 2011). To stimulate innovation development, it is required to examine the influence of the external-driven and internal-driven success factors on innovation (Zirger and Maidique, 1990; Neely et al., 2001; Llorens et al., 2005; Hjalager, 2010). This can be done through recent research areas, namely the innovation theory and the theory of the firm that are resource-based view. The resource-based concepts are able to explain how an organ-
Innovation derivation derives competitive advantage and develop innovative practices externally. This is done by utilising an interactive system of agencies (i.e. public and private institutions) that aim to initiate, produce, diffuse, and use knowledge and technology internally by channelling resources and capabilities into innovation capabilities in order to survive in a dynamic market environment (Mothe and Paquet, 1998; Eisenhardt and Martin, 2000; Hult et al., 2004; Martinez-Roman et al., 2011; Wonglimpiyarat, 2011; Yanadori and Cui, 2013). The study of success factors around firms that condition their innovations, provides extra pathways to understand their innovative capabilities (Martinez-Roman et al., 2011). Innovative behaviour is a complex phenomena that requires various external and internal factors, such as structures for research and development and capabilities for technological advancement in explaining the firm’s innovative development (Neely et al., 2001; Mahemba and De Bruijn, 2003; Laforet and Tann, 2006).

Scholars believe that innovation is vital in a modern environment characterized by hyper-competition (D’Aveni, 1994) and intense and rapid competitive moves necessitate firms to continuously innovate to create new advantages (Dess and Picken, 2000; Tushman and O’Reilly, 1996). Dynamic capabilities are therefore a source of competitive advantage (Eisenhardt and Martin, 2000; Prahalad and Hamel, 1990; Teece et al., 1997), that must be commensurate with the dynamic nature of the environment. According to various scholars, an organization’s tendency to innovate or to adopt innovations is a form of dynamic capability that contributes to competitive advantage (Helfat et al., 2007). For instance, Intel and Rubbermaid used the dynamic innovation capabilities of frequently obstructing competitors by introducing new products and technologies, which facilitated their ability to sustain their “evolutionary fitness” in the market for many years (Helfat et al., 2007). Some dynamic capabilities help organisations for incremental process innovation and lead to experience-related cost reduction (Sinclair et al., 2000) while others, such as drug related innovations, may create and expand new market segments (Bottazzi et al., 2001).

Crossan and Apaydin (2010), proposed that dynamic innovation capabilities exist in managerial levers that enable innovation (Elkins and Keller, 2003; Mumford et al., 2002). Managerial levers can be divided into four distinct categories: missions/goals/strategies; structures and systems; resource allocation; and culture.

3.4.2.1 Organisational mission and strategy

Organizational mission and strategy (Adams et al., 2006) determine the direction for the organisation, while physical and financial resources, organizational struc-
ture, and management and communication systems (Damanpour, 1991) all provide the required support for innovation practices. Organizational learning and knowledge management tools (Crossan et al., 1999) and organizational culture (Pinto and Prescott, 1988; West, 1990) assist in maintaining innovation processes.

An unambiguous innovation strategy (Miller and Friesen, 1982) is a major managerial lever and assists in matching innovation goals with the strategic objectives of the organisation (Tipping and Zeffren, 1995). The strategy types of ‘Prospector’ (Miles and Snow, 1978) and ‘organic’ (Nicholson et al., 1990) have been proposed as critical for innovation.

3.4.2.2 Resource allocation

In terms of resource allocation, the influencing factors include absolute and relative R&D intensity (Parthasarthy and Hammond, 2002), commitment to differentiated funding (White, 2002), annual turnover of resources (Mohr, 1969), and slack resources (Damanpour, 1991; Kanter, 1983; O’Brien, 2003).

For disruptive innovations, the resource allocation process is specifically critical and if not adequate, can cause the failure of the disruptive innovation. The primary inhibitor is structured routines (Nelson and Winter, 1982) such as the key evaluation factor of financial returns (Christensen, 2006) and traditional market research reports. These structured routines constrain the actions of incumbent firms and evaluate emerging disruptive projects by the same criteria applied to existing businesses. Once established, structured routines are difficult to change.

Resource dependence is another inhibitor for disruptive innovation, which means that organisations are locked into businesses in which they have accumulated resources (Christensen, 2006) and tend to invest more in the businesses where firms have sufficient resources. Consequently, organisations often respond to the emergence of competitively threatening technologies and innovations by intensifying their investments to improve the conventional technologies used by their current customers (Christensen and Bower, 1996), thus missing the opportunity for new disruptive innovations. In summary, structured routines (deep-rooted methods to evaluate projects), particularly financial measurements, and resource dependence (investment based on the profile of existing resources) obstruct the development of potential disruptive innovation.

Some scholars have conducted empirical research to solve the difficulties caused by resource allocation in the case of disruptive innovations. One solution is to use strategic buckets to manage sustaining versus disruptive projects independently.
Innovation (Chao and Kavadias, 2007; Hogan, 2005). Another is summarized based on the case study of Good Manufacturing Practice (GMP), that has projects in all the different pipeline phases and manages each phase as a mini-project (Hogan, 2005).

3.4.2.3 Structure and systems

The next managerial levers category is structure and systems that contains organizational complexity and administrative intensity (Damanpour, 1991), specialization and centralization (Damanpour, 1991; Zaltman et al., 1973), formalization (Damanpour, 1991; West et al., 1998), stratification (Kanter, 1983), matrix principles (Staw, 1990), fit between organizational design and type of innovation (Burns and Stalker, 1961), and number of employees (Rogers, 1983) and organisational structure in general.

For disruptive innovations specifically, the aspect of organisational structure covers discussions on (1) the sizes of firm and business units, (2) spin-offs versus ambidextrous organization, and (3) collaboration between incumbent firms and start-ups.

Size of firm and business units investigates the relationship between the number and size of business units and the success of disruptive innovation. Innovation research has largely focused on firm or business unit size as a key success factor for R&D effectiveness (Cohen and Klepper, 1996; Tsai and Wang, 2005) and the recent trend argues that R&D investments are more productive for small than for large firms when introducing new products (Lee and Chen, 2009; Lejarraga and Martinez Ros, 2008). Furthermore, research on disruptive innovations in case studies and surveys of high-tech industries have shown, that the size of the firm is negatively correlated to the success of disruptive innovations (Christensen and Raynor, 2003; DeTienne and Koberg, 2002; Tushman and O’Reilly, 2002).

The implication is that a large corporation should maintain its flexibility by having smaller business units, in order to continue to keep its decision-makers excited and take emerging opportunities seriously. Some authors argue that when more radical innovation is in focus, there is a need for separate structures for innovation that will provide the freedom for the necessary experimentation (eg. O’Connor, 2008; Benner and Tushman, 2003; O’Reilly and Tushman, 2004). Conversely, it has been argued that a separation between mainstream and new-stream work can lead to isolation and resistance to new ideas (Birkinshaw and Gibson, 2004; Moss-Kanter, 2006), and may not take account of the serendipitous ideas that emerge from mainstream work. According to Birkinshaw et al. (2012:1), innovation increasingly is seen as the responsibility of the entire organization, as an “all the time, everywhere” capa-
Innovation

ability. Alänge and Steiber (2013) argue that ultimately the different views about whether to have a separate function for major innovation depends on how innovation is viewed; if it is a continuous or a ‘punctuated equilibrium’. It also depends on whether innovation is focused on how to create new ideas, or whether focus is on nurturing and implementing the ideas, especially if the organizational structure is not open to what is outside of the ordinary, as discussed by O’Connor (2008). Another counter-argument is that redundant overhead expenses can be eliminated when business units are consolidated into much larger entities (Christensen and Raynor, 2003). In addition some researchers argue that organizations must also cope with the demands of architectural interdependence, which often requires larger, more integrated organizations.

On the other hand, one of the main proposals by Christensen to solve the innovator’s dilemma is to set up an ‘autonomous organization’ to develop and commercialize the venture. The key dimensions of autonomy relate to processes and values (unique cost structure) rather than geographical separation or ownership structure (Christensen and Raynor, 2003). Christensen further argues that autonomous organizations are essential due to the fact that most of the incumbent leaders who succeeded at disruption (Chesbrough, 2001; King and Tucci, 2002), had maintained their industry-leading positions by setting up an autonomous business unit and providing it with unrestricted freedom to create a substantially different business model that was appropriate to the situation (Christensen, 2006).

Many scholars argue that it is difficult for a parent organization to equip itself with dual resources, processes and values to manage both sustaining and disruptive innovations, without an isolated autonomous organization. Due to that, Tushman and O’Reilly (2002) proposed the concept of ambidextrous organizations as a solution to manage discontinuous innovations because maintaining a balance between future opportunities and current needs is a challenge for firms; (Tushman and O’Reilly, 1996; Boer and Gertsen, 2003). Mainstream activities (e.g. manufacturing and marketing) contribute to the firm’s current success, with processes building on stability, efficiency and profitability, whereas innovation is unstable, and often requires long-term vision and commitment in order to yield results (Lawson and Samson, 2001). Those organizations that can manage both exploration and exploitation are described as ambidextrous (March, 1991; Tushman and O’Reilly, 1996).

Tushman and O’Reilly (2002) attempted to promote the ability of simultaneously pursuing incremental and radical innovations - which would considerably improve the technical performance and satisfy the existing customers- by citing many successful examples such as HP, J&J, and ABB. However, disruptive innovations, due
to their initial inferior performance, could not attract the same attention from senior managers and existing customers and hence, the potentially valuable contributions of the authors may not be applicable in the case of disruptive innovations.

Some other researchers found that Open Innovation Theory may be applied to managing disruptive innovations (Chesbrough and Crowther, 2006; Paap and Katz, 2004). New start-ups continually possess innovative and potentially disruptive technology strengths, however they typically lack the complementary assets that are the core strength of incumbent leaders (Rothaermel, 2001). Some researchers have empirically discovered that spin-offs, alliances, market transactions and acquisitions are comparatively optimal for corresponding points at different stages of disruptive innovation (Claude-Gaudillat and Quelin, 2006). The empirical multiple case studies by Macher and Richman (2004) found that IBM, Kodak and HP have adopted either form or a combination of different forms of collaborations to create disruptive innovations. In addition to research on different forms of open innovation, collaborator attractiveness has also been studied. A survey by Rothaermel (2002) found that a start-up’s new product development, economies of scope, public ownership and geographic location in a regional technology cluster are positively associated with the start-up’s attractiveness as an alliance partner to create potentially disruptive innovation.

3.4.2.4 Organisational culture

Organisational culture is another critical factor that enables innovation as a process. Innovative cultures are created by leaders through having a clearly stated, attainable, and valuable shared vision (Pinto and Prescott, 1988; West, 1990), which promotes autonomy (Amabile, 1998; Zien and Buckler, 1997), calculated risk taking (West, 1990), and motivation (Miller and Friesen, 1982).

Culture is an effective way of controlling and co-ordinating people without elaborate and rigid formal control systems (Tushman and O’Reilly, 2002). However, culture can occasionally result in the failure of innovation. When great changes such as disruptive innovation occur, case studies have revealed that the organizational culture generates cultural inertia that is significantly challenging to overcome directly and is therefore a key reason why managers often fail to introduce timely and substantial change, even when they know that it is required (Christensen and Raynor, 2003; Henderson, 2006; Tushman and O’Reilly, 2002).

Consequently, it is crucial for incumbents to prepare for and institute organizational change and unlearn deeply entrenched values at the advent of potential disruptive
innovation. Nevertheless, some integral elements of culture, such as entrepreneurship, risk-taking, flexibility and creativity, should be preserved and valued in order to develop disruptive innovations (Govindarajan and Kopalle, 2006; Murase, 2003).

Learning environment is also critical to the innovation culture. A learning environment is cultivated by leaders through provision of support for experimentation (Damanpour, 1991; King et al., 1992; West and Anderson, 1992); being tolerant of failed ideas (Madjar et al., 2002); adopting risk-taking norms (King et al., 1992; West and Anderson, 1992); supporting learning and development of employees; and fostering the acceptance of diversity within the group (Crossan and Hulland, 2002). These factors are also critical parts of the organisational innovation culture.

Together, these four managerial levers (mission, goals, and strategy; structure and systems; resource allocation; and organizational culture) enable core innovation processes.

3.4.3 Business processes

Business Processes supporting innovation is a meta-construct that merges process-level variables. This meta-construct is debatably the most developed in the literature within the framework of process theory, and studies how organizational processes convert inputs into outputs.

The word ‘process’ has an extensive variety of meanings and therefore I begin with clarifying its application in this thesis. According to Van de Ven and Poole (1995), in management literature the term ‘process’ refers to: (1) the underlying logic that describes a causal relationship between independent and dependent variables in a variance theory; (2) a category of concepts of organizational actions, such as rates of communications, work flows, decision making techniques, or methods for strategy creation; and (3) the progression (i.e. the order and sequence) of events in an organizational entity’s existence over time. In this section we use the second interpretation of the word when referring to Business Processes.

The process approach has been used extensively in various areas of social science and examples include labours process theory (Knights and Willmott, 1990), process theories of human behaviour (motivational theories: Adams, 1963; 1965; Kahler, 1975; Locke, 1968; 2001; Vroom, 1964) and cognition (information processing theory: Miller, 1956).

Typically process theories are established based on the belief that similar inputs transformed by similar processes will lead to similar outcomes; provided certain con-
stant conditions that are necessary for the outcome to be reached, are met. Therefore, a process level explanation uncovers the generative mechanisms that cause the observed events to occur, and the specific circumstances or contingencies when these causal mechanisms operate (Harré and Madden, 1975; Tsoukas, 1989).

According to Van de Ven and Poole (1995), in process theory, usual patterns of events are core theoretical constructs. In the context of innovation, these core processes are normally divided into three parts: the front end of innovation, the new product/service development process, and commercialization. The front end (often referred to as the fuzzy front end) is the summation of all activities that precede the well-structured New Product Development (NPD). Front end includes all the time spent on the idea in addition to the activities enforcing it. Hence, this phase starts with the first impulse and/or opportunity for a new product or a new service and continues up to the decisions concerning implementation and the start of development of the new product and/or service (Reil, 2011), (Herstatt et al., 2005). The front-end is challenging as it requires to combine adequate space for creativity and freedom of ideation, with systemised activities to enhance efficiency (Herrmann et al., 2007).

Various innovation models have been proposed in literature to identify the core constructs in an innovation process. Hobday (2005) reviewed five generations of innovation models developed from the 1950s to the 1990s (technology push, marketing pull, coupling models, integrated models, and networking models) and confirmed Mahdi’s (2002) finding that even the latest innovation models failed to consistently explain findings across and even within sectors. The author argues that intra-sector disparities are due to the path-dependent and iterative nature of the innovation process, and therefore an appropriate model should adopt an evolutionary approach and allow iteration. These concerns were subsequently addressed by Van de Ven et al. (2007), who proposed to view innovation as a non-linear dynamic system that comprises of a cycle of divergent and convergent activities, which may be repeated over time and at different organizational levels. This innovation cycle is enabled by resource investments and organizational structure; and the boundaries of the journey are drawn by external institutional rules and internal focus.

Table 3.2. shows the models of innovation that have evolved overtime from a simple linear model to an integrated and networked model to a total innovation management model Hobday (2005).

The innovation model has been changing from a manufacturer-centred innovation paradigm to a collaborative user-centred innovation paradigm (Hippel et al., 2001;
Von Hippel, 2006). Many scholars including Selden and MacMillan (2006) have suggested the use of a customer-centric innovation model when innovation is intended to deliver results that meet or exceed market expectations.

In our research we are aiming to create a model for enabling data analytics driven innovation in large organisations; with an emphasis on the front end of the innovation process - the identification and selection of an idea with an innovation potential. An increasing number of studies highlight the importance of the front-end of innovation for the overall success of innovations (Broring, 2005), (Cooper, 2011), (Kim et al., 2002), (Stevens et al., 2003) and believe that it is the root of success for firms involved with discontinuous innovations. This is in line with our empirical findings in section 2.3. where the front end of innovation was shown to be critical to the success of data analytics projects resulting in innovation. According to the Front end model developed by Reid and de Brentani (2004), this is because discontinuous innovations are typically not the result of explicit and structured organizational processes, as is the case for incremental innovations, but rather of complex “bottom-up” phenomena, that occur during this very early phase of the innovation process.

In their research, Reid and de Brentani (2004) specifically focus on discontinuous innovations that, as defined by Garcia and Calantone (2002), entail “radically new” innovations (i.e., those requiring changes in both existing technology and marketing infrastructure) and “really new” innovations (i.e. those involving either technology or marketing discontinuities). This is a critical area of focus, because this type of innovation has the highest level of uncertainty during the front-end and in addition the development of products resulting from such innovations involves the greatest lack of understanding and the fewest strategies for effective management.

As Kim et al. (2002) argues, the effective management of the early phase of the innovation process is the origin for innovative ideas for sustainable competitive ad-
Innovation

vantages. This influence of the front-end on new product development has been
demonstrated by a number of empirical studies (Hermann et al., 2007), (Verworn,
2006), (Verworn et al., 2008), (Stockstorm et al., 2008) and Table 3.3. reviews the
key results of these studies.

<table>
<thead>
<tr>
<th>Object of investigation</th>
<th>Results</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>144 German measurement and control firms</td>
<td>Companies that systematically reduce market and technological uncertainties during the front-end of innovation, repeatedly outperform other and are among more successful innovators.</td>
<td>(Verworn, 2006)</td>
</tr>
<tr>
<td>497 New Product Development (NPD) projects from Japanese mechanical and electrical engineering firms</td>
<td>Key driver of project success is the intensity of planning prior to the start of development: relationship between front-end factors and project success.</td>
<td>(Verworn et al., 2008)</td>
</tr>
<tr>
<td>475 Research and Development projects in Japanese electrical and mechanical engineering companies</td>
<td>Planning intensity during the early phase of innovation is linked to the project success</td>
<td>(Stockstorm et al., 2008)</td>
</tr>
</tbody>
</table>

Table 3.3: Studies confirming the impact of the front end on NPD

The studies shown in table 3.3. indicate that early reduction of technical and market
uncertainty, involvement of all project members and top management, and inter-
disciplinary teamwork and communication, can affect the success of the front end significantly.

Verworn and Herrstatt (2008) believe that because the degree of uncertainty is at its height in the front-end of innovation processes, flexibility has the utmost priority. They propose that the management of the front-end is required to be adapted to the level of uncertainty for the various types of innovations and therefore innovation strategies and process models need to reflect the respective market and technology uncertainties (Verworn et al., 2007). The matrix in Figure 3.3. below depicts the four types of innovation, each representing a distinct degree of market and technology uncertainty (Verworn et al., 2007).

Based on this uncertainty matrix it is possible to depict the relationship between the type of innovation and the required degree of formalisation and therefore different innovation management models are applicable to these different types of innovations. Since the front-end of innovation (FEI) -where opportunities and ideas are identified and generated- is a core part of this PhD, I will take a deeper look into the FEI process models, with the aim to provide a basis for system and project process design decisions within the context of CDA.
3.4.3.1 Structured problems

Structured and process-oriented activities are effective for incremental innovations with low market and technological uncertainty, and can contribute to an efficient implementation. In such innovations, the planning can be conducted with a high degree of accuracy and consistency, because the technologies and market conditions are mainly known. In addition, external forecasting techniques such as customer surveys can result in reliable predictions (Lynn et al., 1998).

For incremental innovation, one of the most popular models in industry and widespread among professionals (Cooper et al., 1990), (Cooper et al., 1991), (Rundquist and Chibba, 2002), (Whitelet et al. 1998) is the stage-gate process by Cooper (Cooper et al., 2011). The model has been implemented in a significant number of organizations such as 3M, Procter & Gamble or Hewlett Packard (Verworn et al., 2007) since the late 1980s to organise their development processes (Cooper, 1988), and make product development more efficient by reducing uncertainties as early as possible (Wheelwright and Clark, 1992). While the stage-gate models have aided firms in shortening their lead times, research has discovered that a high degree of formalization may be disadvantageous to innovation (Benner and Tushman, 2002, ref), and it has been argued that established New Product Development (NPD) processes do not result in innovative products (Veryzer, 1998; Leifer et al., 2001; McDermott and O’Connor, 2002). This is because normally in complex development processes, there is little room for experimentation and learning (Engvall, 2003), which means the ability to absorb new knowledge decreases (Christensen, 1997).
Many scholars argue that while this type of uncertainty-reducing processes may be suitable for developing pre-defined products, it will not be adequate for products or solutions that are radically different (Veryzer, 1998; Eisenhart and Tabrizi, 1995; McDermott and O’Connor, 2002; Engvall, 2003). Engwall (2003) points out that when the ideas for what to develop exist from the beginning there is little room for truly innovative concepts. Ideas and required knowledge are considered as inputs, and not as evolving throughout the project, which therefore results in minute learning during the execution of the project.

The stage gate process

In the stage gate model, the innovation process is divided into individual, sequentially proceeding phases called “stages” and the various stages are multifunctional. Subsequent to each phase, there is a “gate”, where a decision will be made in regards to the continuation or termination of the project and/or if the conditions to proceed to the next phase are met. The gates act as a check point for whether a particular phase was conducted adequately and if the necessary deliverables have been accomplished (Verworn et al., 2007).

Figure 3.4. illustrates a stage-gate model for the early phase of innovation. Initially the ideas are generated using internal or external sources and during the initial screening, a decision is made in regards to the allocation of resources to develop the idea (integrating both the market and technology perspectives, which will be processed in parallel). A decision is subsequently made at the second gate, based on these engrossed information, in regards to whether the idea will be developed into a concept. If a positive decision is obtained, the collected information will result in a product concept and after the development of a technical concept to implement the idea, the market acceptance will be tested through market studies. Depending of the results of the concept tests, a decision will be made in regards to the implementation of the concept and further allocation of resources. A cross-functional team is involved in all the individual gate decisions (Verworn et al., 2007).

The main advantages of a stage-gate process lie in the fact that a common understanding about the steps of the innovation process can be achieved. It provides clear objectives against which projects can be assessed at each gate and after each phase, a review of the implementation takes place. Consequently, a previously ad-hoc approach of development is systematised to increase the efficiency and effectiveness of each stage (Verworn and Herstatt, 2007).

Conversely, as previously discussed, the main criticism of the stage-gate process, similar to other processes that are suitable for incremental innovation, lies in its
sequential design and its lack of flexibility. In addition, the first stage-gate models described by Cooper paid significantly less attention to the early stages, in particular the idea phase. Later, in order to integrate non-directional fundamental research, Cooper and colleagues introduced an additional process chain, the “discovery stage” for technical developments, which should take into account the experimental nature of technology-induced innovations. However, this approach also failed in the detailed description of the ideation, because the phases are extremely roughly defined. The activities are much diversified and again there is no concrete explanation for the generation of ideas.

Evidently, the stage gate model of front end innovation and rigorous models similar to it that do not allow for experimentation, are not suitable for the design of the CDA framework.

3.4.3.2 Unstructured problems

In situations where the market uncertainty is low and the technological uncertainty is high, or vice versa, the focus should be on activities that exploit the existing knowledge and reduce the residual risk. In such cases, separating the ideation process to strictly sequential phases will not meet the requirements of reducing and minimising technical or market uncertainty. According to Verworn et al. (2007) in these situations, a learning-based strategy and an iterative procedure is required. Eisenhardt and Tabrizi (1995) believed, in uncertain conditions, an ‘experimental model’ can be a superior fit, allowing for improvisation and flexibility, and where learning is
accelerated through iterations and testing, in combination with strong motivation and leadership.

The most extreme case of innovation is a radical innovation that seeks both new markets and technologies. For these innovations, all areas and functions must proceed gradually through extensive processes of learning and experience. For this purpose, the process must have the necessary openness to guarantee iterations and to make the integration of feedback possible at the right time (Verworn et al., 2007).

In this PhD thesis, the focus is on unstructured problems and therefore a model that is flexible and allows for learning through iterations and testing is required. In the next two sections we discuss two of the most important models presented in literature for managing the front end of discontinuous innovations.

**Probe and learn process**

This process highlights the aspect of learning-based strategies concerning the front-end of innovations. Based on the examination of four successful radical innovations, Lynn et al. derived the Probe and Learn Process to fulfil the specific requirements of high technical risk and/or market uncertainty (Lynn et al., 1996).

The Probe and Learn Process is specifically designed to reduce uncertainty during the early phases of innovation and corresponds to the iterative procedure and learning-based strategy that Verworn and Herstatt recommend for radical innovations (Verworn and Herstatt, 2007).

For radical innovations, at the time of market launch, the design and the potential customers are unknown and consequently Lynn et al. propose an iterative procedure: Early versions of products will be introduced to test markets, modified based on the obtained learnings and experiences, and hence re-tested in the market. These iterations will be repeated until all essential information have been gathered and verified.

The first step (“probe”) has the character of an experiment where an initial product version is introduced to a plausible preliminary market. For instance, General Electric tested a breast scanner to enter the Computer Axial Tomography (CT) business in the mid-1970s and a whole body scanner was subsequently developed based on the lessons learned from this test. The aim of the experiments should be to obtain adequate and targeted information, enabling the next iteration of the product or service. The innovation process for developing a whole-body scanner is shown in Figure 3.5. (Lynn et al., 1996).
The iterative Probe and Learn Process is different from the conventional sequential stage-gate processes, in that the emphasis is on learning-based creation of new knowledge and not on process efficiency. None of the products described by Lynn et al., would have passed one of the gates of a sequential process during the early phases of innovation. Hence, this model is not applicable in the field of incremental innovations, and instead is beneficial in areas of high uncertainty, that can only be reduced through learning, which include radical innovations, technical innovations and market innovations (Verworn and Herstatt, 2007).

**The new concept development model**

In 1998, Khurana and Rosenthal published the first comprehensive study of the front end of innovation based on case studies of 10 incremental and 2 radical projects.
They found that successful organizations follow a holistic approach, one that addresses the front end within a broader organizational context, and that success depends on both organizational attributes and project-specific activities.

In 2001, an Industrial Research Institute (IRI) ROR project team extended Khurana and Rosenthal’s work by creating a holistic framework for the front end, called the New Concept Development (NCD) model (Koen et al., 2001). That work also introduced the term “front end of innovation,” intended to replace the more expressionistic term “fuzzy front end,” coined by Reinertsen (1985), with its implications that the front end is mysterious, lacks accountability, and cannot be managed. Koen and his colleagues try to explain the front-end with the objective to design a model that represents the character of this phase rather than developing a reference process. As a continuous progress of the holistic perspective from Khurana and Rosenthal, the NCD model includes in addition to development activities also internal and external factors. This theoretical construct provides a common language and definition of the key components of the front-end of innovation (Koen et al., 2001) and (Koen et al., 2002).

As seen in the image below, the NCD model divides the front end into three distinct area: the engine, the wheel and the rim.

The engine, at the centre of the model, provides power to the front end of innovation. The engine consists of two separate segments – organizational attributes and teams and collaboration. The majority of studies in the literature that evaluated radical innovation focused on organizational challenges rather than on activities (Koen et al., 2014). The wheel, the inner part of the model, comprises the five activity elements of the front end: opportunity identification, opportunity analysis, idea generation, idea selection, and concept definition. The third element, the rim, includes the environmental factors that influence the engine and shape the five activity elements. These include the company’s organizational capabilities, competitor threats, customer and worldwide trends, regulatory changes, and the depth and strength of enabling sciences and technology. Below a summary of these three major parts are given:

1. The influencing factors bundle the peripheral environment of the process. These factors are on the one hand internal factors such as the organisational capabilities, business strategy, enabling science and technologies, and on the other hand also external factors like the outside world (government policy, environmental regulations, laws concerning patents and socioeconomic trends), distribution channels, customers and competitors. The influencing factors are
sources of new ideas and affect the entire innovation process, including the fuzzy front-end as well as the NPD and commercialisation.

2. The core of the model is the engine. It includes the leadership and corporate culture and drives the five front-end elements.

3. The five controllable front-end elements consist of the following activities (no sequential order):

   - Opportunity Identification concerns the identification of product or market opportunities, that the company wants to pursue and which are driven by the company’s objectives

   - Additional information is collected during the Opportunity Analysis to assess the value of the opportunity. So it is possible to translate the identified opportunity into specific business and technology opportunities. The extent of the effort for the analysis depends on the information needed to reduce uncertainties. Typical questions are: How attractive is the opportunity? What size has the future development effort? Does the
opportunity fit with the corporate strategy and culture? How high is the decision makers’ risk tolerance?

- The element of Idea Generation and Enrichment represents the birth, development and maturation of an idea. Through the integration of customers or users and other external stakeholders, like collaborations with other companies and institutions, the opportunity is evolutionarily modified to a concrete idea. Also cross-functional teams enhance the idea generation. This element of Idea Genesis can also be encouraged from the outside, for example through new materials available on the market or random test result in the laboratory. The result of this part of the NCD is usually a detailed idea description or a product concept.

- The output of the idea generation is the subject of the next element, called Idea Selection. Here a first evaluation of the idea happens. As the level of information at this stage has a still great deficit, and financial details are usually very roughly estimated, Koen et al. show the need for a multidimensional evaluation approach. Possible assessment criteria are investments, risks, competition, existing competences and the product benefit.

- The last element of the NCD model is the Concept Definition. The selected ideas have to be concretised by the development of a business case, which includes estimates for investment in the business or technology. The formality of the business case depends on several factors, like the nature of the opportunity, level of resources, the organisational requirements to proceed to the NPD and the corporate culture. With the development of the business plan and/or a formal project proposal the final deliverable has been completed, and the idea can be transferred from the NCD to the NPD process.

A key insight of the NCD is the identification of two starting points for an innovation initiative. These are represented by the two arrows pointing into the model, that indicate projects begin at either opportunity identification or idea generation and enrichment. There is also an exiting arrow representing how concepts leave the model and enter the new product development (NPD) or technology Stage-Gate process (or any other type of innovation development).

Also, the inner parts of the NCD were designed as elements rather than processes. This contains the explicit reference to the iterative nature of the described activities, also graphically represented through the circular shape. Ideas are expected to flow, circulate and iterate between and among all of the five front-end elements. This is
Innovation

also, in contrast to linear, staged-and-gated processes (Ajamian and Koen, 2002). Furthermore, the NCD takes into account the influence of the internal and external environment to specific activities. Finally, intensity of the activities relies on the content of the opportunities, like the degree of innovation, and on the corporate culture (Globocnik, 2011).

Koen et al. (2014) argue that the language of the NCD model provides a vocabulary for understanding the activities that occur in the front end. In the NCD model, an “opportunity” is defined as a business or technology gap that exists between the current situation and an envisioned future. For example, a food company may identify a growing need for low-fat products arising from increased consumer interest in the healthful eating but have no ideas and few products that can be sold in this segment. An “idea” is defined as the most embryonic form of a new product or service, such as an idea for a new food category – for instance, nonfat potato chips. The food company’s research group may envision several ideas for molecules that provide the same flavor like fat molecules. A “concept” refers to a well-defined product or service with principal features and customer benefits clearly identified. For the example of the food company, the outcome of the R&D program arising from the idea could be a scientific program to create nonfat molecules, resulting in a new product that has the same taste as the original product but contains no fat.

Below is a summary of the above terms with their definitions:

1. Opportunity: A business or technology gap, that a company or individual realises, that exists between the current situation and an envisioned future in order to capture competitive advantage, respond to a threat, solve a problem, or ameliorate a difficulty.

2. Idea: The most embryonic form of a new product or service [or any other form of potential innovation]. It often consists of a high-level view of the solution envisioned for the problem identified by the opportunity.

3. Concept: Has a well-defined form, including both a written and visual description, that includes its primary features and customer benefits combined with a broad understanding of the technology needed.

Koen et al. (2014) state that these expanded definitions provide a vocabulary to explain the activities that occur in the front end. Without them, opportunity identification and ideation are frequently confused. Opportunity identification refers to the process of identifying new markets with unmet needs and emerging trends, while ideation is concerned with finding embryonic solutions to the problems identified in
Innovation the opportunity space. According to Keon et al. the reason this matters is that many organisations begin a new innovation initiative by asking for new ideas. By doing so they are making the implicit assumption that the biggest growth and profit will come from unmet needs in the existing market (an opportunity space the company is familiar with). Instead, Koen et al. however believe that organisations should start by identifying the largest opportunities, then develop ideas that can fill that opportunity space. For example, when IBM was looking for new growth, the company explored the healthcare opportunity space, although it was at the time a new opportunity for IBM, because it had large growth potential (Garvin and Levesque, 2005).

Koen et al. (2014) conducted a study based on data collected from 197 large US based companies, and analyzed the elements important for success in the front end of innovation. They developed eight constructs for activities associated with radical innovations (they were also aligned with most of the activity elements in NCD model):

- **Opportunity Identification (Trends):** The degree to which the business evaluates economic, demographic, consumer, and cultural trends as well as regulatory shifts.

- **Opportunity Identification (Disruption):** The degree to which the business evaluates opportunities that are potentially disruptive to their current businesses.

- **Opportunity Analysis (Tools):** The degree to which the business unit uses tools such as technology roadmapping, scenario planning, and product generation mapping in their radical innovation activities.

- **Opportunity Analysis (Selection):** The degree to which the business unit evaluates the probability of market and technical success and does a competitive analysis and assessment of the advantage.

- **Idea Generation (Ethnography):** The degree to which the business uses ethnographic methodology to identify new ideas in order to understand unarticulated reasons customers make product choices.

- **Idea Generation (Technology):** The degree to which the business unit obtains new ideas from new technologies or technology-driven inventions.

- **Concept Definition:** The degree to which the business unit determines the feasibility of the radical innovation concept from market, customer, and com-
Koen et al. (2014) found that three factors were significantly related to front-end performance: opportunity generation focused on identifying trends, opportunity generation focused on identifying disruptive businesses, and idea generation focused on new technology. The first two of these factors capture a process for identifying radical innovation opportunities that is fundamentally different from that required for incremental innovation.

3.4.4 Customers

Determining customer orientation under disruptive changes is another major stream of literature that attempts to seek solutions from the customer’s perspective. It has been argued that insights about user needs are a superior starting point to technological development or visions within the firm (Veryzer and Borja deMozota, 2005; Jansen and Dankbaar, 2008). On the other hand, some researchers have raised concerns that addressing user needs inescapably leads to incremental innovation (Hamel and Pralahad, 1994; Verganti, 2008; Christensen, 1997; Govindarajan et al., 2011).

In 2002 Koen et al. conducted a study that showed ethnography idea generation is not significantly related to positive front end performance. Similarly, in their study Govindarajan et al. (2011) presented that an orientation to mainstream customers was significantly and negatively related with radical innovation and that an orientation to mainstream customers was significantly and negatively related with radical innovation and that an orientation to small, emerging customer segments was positively related to radical innovation success. The importance of carefully finding the emerging market and deeply understanding the customers’ latent needs cannot be emphasized more, because a firm’s disabilities in finding new markets for new technologies may be its most serious innovation handicap (Christensen and Bower, 1996).

Danneels (2002) proposed that a second-order marketing competence is the ability to add new customers to address new markets. When established companies are not blind-sided by the development of new technological capabilities, failing to link the development of such technological advances to changes in the marketplace and consumer needs or market conditions is likely. The main way to avoid the negative effects of disruptive innovations is to focus on what is happening with customer and operational needs. Govindarajan and Kopalle found that the higher an SBU’s emerg-
Innovation

ing customer orientation, the more disruptive the innovation it developed would be (Govindarajan and Kopalle, 2004).

On the other hand, other scholars argue that mainstream customer orientation and emerging customer orientation are independent of each other and that firms can develop both orientations simultaneously (Baker and Sinkula, 2005; Narver et al., 2004; Slater and Mohr, 2006). They argue that customer-oriented firm can serve current customers and also remain vigilant of non-consumption in emerging markets (Chandy and Tellis, 1998; Day, 1999; Slater and Mohr, 2006).

Handerson (2006) proposed that incumbent firms failed to respond to disruptive innovations as a result of being ill-equipped in being able to acquire the competencies that were required to respond to disruptive innovation appropriately and not because they focused too much on existing customers and high margin opportunities. In fact, in some cases, incumbent firms identified the needs of the low-end or emerging customers but lacked the market-related competencies to respond quickly to disruptive innovation.

While disruptive innovations typically start from the bottom of the market focusing on emerging customers, giving incumbents time to react, big-bang disruptions provide no such notice. According to Dennings (2014) big bang disruption is a symptom of an even broader management challenge that results from corporations no longer being at the centre of the economic universe. Technology has enabled significant amount of new possibilities for producers and at the same time it has provided customers with immense power by enabling them to have instant access to reliable information and an ability to communicate amongst each other. Consequently, the power in the marketplace has shifted from seller to buyer.

According to Dennings (2014), big bang disruption is just one of the more significant symptoms of this broader and deeper economic phase change: the emergence of the creative economy, where continuous transformational innovation is occurring. In this era, it is no longer possible for corporations to prosper by efficiently producing products and services and pushing them at passive consumers (as they had in the 20th century). Coping with this new context lies on understanding, anticipating, and meeting the needs, wants and whims of customers that are well informed, empowered and interactive. Therefore corporations must learn to delight customers through continuous, disciplined, transformational innovation (Dennings, 2014).

A variety of methods for incorporating user knowledge have been proposed, in particular those that stress the importance of addressing tacit and future needs of current and future customers rather than relying on what users actually say (e.g.
Von Hippel, 2009; Leonard and Rayport, 1997; Goffin et al., 2010; Kristensson et al., 2004; Narver et al., 2004). For example, ethnographic and empathic approaches have been used to study customers in their contexts (e.g. Burchill and Fine, 1997; Leonard and Rayport, 1997), and a variety of co-creation approaches have been discussed involving users in, for example idea generation and selection (Öberg, 2010), or a “probe and learn” approach where users are exposed to and provide feedback on immature prototypes (such as beta testing) (Lynn et al., 1996; Cole, 2002).

These methods are important as the creation of ideas cannot be left to chance during a structured ideation process. Thus, it is important to identify the most promising ideation activities that can stimulate creativity actively. In a study published by Cooper and Edgett (2008), they tried to find an answer to the question: “Ideation for Product Innovation: What are the best methods?” Their study looked at 18 different ideation methods with the objective to determine how extensively each ideation method is used (the popularity of the method) as well as to gauge management’s perception of the effectiveness of the method in generating excellent, high-value new product ideas. A total of 160 companies took part in the study conducted in 2007 (Cooper et al., 2008)

Figure 3.7. below presents the popularity and effectiveness of each of the 18 methods in the magic ideation quadrant diagram proposed by Cooper (2008). The popularity is measured by the percentage of firms that extensively use each method (usage was captured on a 0-10 scale; “extensive users” are those that checked the top third of this 10-point usage scale). Rated effectiveness of each method is presented as a 0-10 scale, but only for users of that method. Ideation methods that are both popular and effective are in the upper right quadrant, approaches that are not too popular and rated ineffective are in the lower left quadrant (Cooper et al., 2008).

The magic ideation quadrant diagram in Figure gives a good overview of the popularity and effectiveness of voice-of-customer methods. One of the most important and popular methodologies for customer centric innovation is Design Thinking. As Brown (2009) describes it, Design thinking is “a human-centered approach to innovation that draws from the designer’s toolkit to integrate the needs of people, the possibilities of technology, and the requirements for business success.”

3.4.4.1 Design thinking

In innovation research there is a growing interest in how design is linked to radical, game-changing and breakthrough innovation (Verganti, 2008; Bruce and Cooper, 2003; Bruce and Bessant, 2002). In 2013 Jahnke performed an experimental study
of professional designers intervening in product development in five Swedish ‘non-designerly’ firms over two years, to investigate the influence of the designers on the the front end of innovation in these firms. Adopting a hermeneutic perspective, Jahnke (2013) shows that the designers helped the firms’ representatives to reflect critically upon their preconceptions of their products, users and the firms. According to Jahnke (2013), this contributed to expanding their horizons of understanding, which provided new meaning-spaces for innovation. He found also that the process was facilitated by collective engagement in the activities of ‘aesthetic deliberation’, in which designers and firm representatives use visual tools such as sketches or simple models built from for example foam and paper, not to create concepts but to explore and learn together. Öberg (2012) and Verganti and Öberg (2013) also explore the role of design in innovation from a hermeneutic perspective, suggesting that the creation of meaning can be seen as a special form of radical innovation. In an extensive ethnographic study, Stigliani and Ravasi (2012) examine designers in a design firm, and find that visualization methods are crucial when designers engage in future-oriented group processes. They conclude that this use of material practices in prospective sense making could be applied to all kinds of generative work in firms, particularly innovation and strategy.

In line with this increased interest in design in managerial debates, the concept design thinking (DT) in management emerged in the early 2000s, originating in the practical experience of managers in the design firm IDEO (Kelley and Littman, 2001; Brown, 2009), and management scholars who had collaborated with or observed the work of designers (Martin, 2009; Boland and Collopy, 2004). DT is described
as a multidisciplinary human-centered innovation approach inspired by the ways designers think and work (Brown, 2008; Johansson-Sköldberg et al., 2013; Kimbell, 2011). While the significance of the professional designer is emphasized by both innovation and design researchers, the core idea of DT is that any firm can be inspired by designers, thus stepping away from the active role of professional designer in innovation work (Brown and Katz, 2011; Brown, 2009). DT is now practiced in many large organizations, including SAP, P&G, Intuit and Kaiser Permanente (Martin, 2011; Holloway, 2009; Lafley and Charan, 2008; McCreary, 2003). It is applied also to many areas such as products, services, or social innovation (Brown, 2009).

DT has been described as “one of the hottest trends in business” (Liedtka and Ogilvie, 2011), and is advocated in practitioner books, business press (e.g. Kelley and Littman, 2001; Brown, 2008; Brown, 2009; Martin, 2011; McCreary, 2010; Liedtka and Ogilvie, 2011) and Design Management Institute events and conferences on the topic (Lockwood, 2009) and published journal articles. However, due to its lack of a theoretical foundation, the concept of DT is dismissed by some as a fad (Johansson-Sköldberg et al., 2013; Jahnke, 2013; Rylander, 2009).

In tracing the roots of design thinking, Johansson-Sköldberg et al. (2013) suggest three principal origins: 1) IDEO - focusing on the way in which the design firm works with clients, often with a hands-on approach (Brown, 2009; Kelley, 2008), 2) Roger Martin7 - focusing on innovative ways of thinking and the skills necessary for managerial success (Dunne and Martin, 2006; Martin, 2009), and 3) Management theory (Boland and Collopy, 2004). The latter two focus on DT as more of a cognitive process or resource. DT is often being described as a creative, subjective and emotional alternative to the analytical logic characterizing many large firms (Brown, 2008; Brown, 2009; Rylander, 2009), although Martin (2009) refers to DT as supplying the balance between analytical and intuitive thinking, stressing that neither logic is sufficient on its own. Most proponents of DT describe how it is influenced by how designers think and work, but focus on how it takes account also of aspects such as feasibility and viability, and that one of its core aspects is the creativity that emerges from the tensions among these various constraints (Brown, 2008).

The label of DT is causing some confusion since the study of how professional designers and architects work and what though processes are going on, goes under the same name. Johansson-Sköldberg et al. (2013), argue that research on DT can be categorized as: (1) ‘designerly thinking’ that pertains to the design research tradition of studying designers that goes be traced back to the 1960s, and (2) “de-
sign thinking” that is related to the emerging managerial concept. According to Johansson-Sköldberg et al. (2013), the academic stream and the recent hyped discourse are different in nature, with the first focusing on professionally educated designers and teams, while the second often discusses multidisciplinary teams of industry trained ‘design thinkers’, performing outside of the realm of traditionally trained designers. They further describe that these two discourses are totally disconnected and that there are few, if any, cross-references between the fields (ibid). It is, however, common for researchers to not make this distinction; for example Hobday et al. (2011) who also refer to Verganti and Hatchuel as design thinkers. In this thesis, the distinction made by Johansson-Sköldberg et al. (2013) is adopted and DT refers to current managerial discourse.

In recent years, DT has started to generate interest among scholars, both in innovation and management research (Seidel and Fixson, 2013; Ward et al., 2009; Wyman et al., 2012; Plattner et al., 2012; Chang et al., 2012), as well as in design research (Johansson-Sköldberg et al., 2013; Kimbell, 2011). However, there are still a paucity of peer-reviewed articles on this topic, and attempts to review the literature of design thinking mainly relies on books and articles in the business press (e.g. Brown, 2008; Dunne and Martin, 2006; Kelley and Littman, 2001), as well as conference papers written by scholars (Hassi and Laakso, 2011).

### 3.4.4.2 Understanding design thinking as a concept

A hotly debated topic is how to understand the concept of DT in theory and in practice. As Johansson-Sköldberg et al. (2013) and Kimbell (2011) note, DT is a rather loose term that can have several different meanings. Descriptions of DT range from a prescriptive process where multidisciplinary teams use a user-oriented approach to come up with relevant solutions to ‘wicked problems’ (to use the vocabulary of design research) to a set of cognitive characteristics that managers can learn from designers (Kimbell, 2011; 2012; Johansson-Sköldberg et al., 2013). Summarizing the various practitioner-based descriptions of DT, Jahnke (2013) states that DT is often understood as a problem-solving approach to innovation, in line with Herbert Simon’s perception of design.

Roger Martin (2009) conceives DT as the ability of professional designers to switch between abductive, inductive and deductive ways of reasoning. He states that if managers were to adhere to this viewpoint, they would not only choose between given alternatives, but also come up with entirely new solutions. According to Martin (Dunne and Martin, 2006), DT in practice could help managers to cope with classical challenges such as balancing between exploration and exploitation. Tim Brown of
IDEO defines DT as a discipline that uses “the designer’s sensibility and methods to match people’s needs with what is technologically feasible and what a viable business strategy can convert into customer value and market opportunity” (Brown, 2008, p.86). Similarly, IDEO defines DT as “a human-centered approach to innovation that draws from the designer’s toolkit to integrate the needs of people, the possibilities of technology, and the requirements for business success” (www.ideo.com).

Thus, there is no single common perspective on or definition of DT. Researchers have called for an ‘epistemological attention to the discourses’, strongly rejecting the idea of a single definition of DT (Kimbell, 2011; Johansson-Sköldberg et al., 2013; Hassi and Laakso, 2011).

Johansson-Sköldberg et al. (2013, p.3) claim that:

“As social constructionists we regard an approach that begins with the question, ‘What is design thinking?’ as an essentialist trap. We do not believe that there is a unique meaning of ‘design thinking’, and accordingly we should not look for such a one. Instead, we look for where and how the concept is used in different situations, both theoretical and practical, and what meaning is given into the concept.”

Some researchers conceptualize the content of DT. Based on a critical review of the literature on DT (designerly and managerial), Kimbell (2011) characterizes DT as: 1) a cognitive style of individual designers involved in problem solving, 2) a general theory of design as a field or discipline focused on solving wicked problems, 3) an organizational resource for businesses and other organizations. One approach is to depict common elements from the DT literature. Based on a literature review of DT in the managerial literature, Hassi and Laakso (2011) describe DT within a three-dimensional framework of practices (ways of working), thinking styles (cognitive styles and ways of processing information), and mentalities (the mental attitudes of individuals and the organizational culture). Another attempt to identify common elements of DT is provided in the literature-based study by Seidel and Fixson (2013) who propose three broad methods: 1) need finding, encompassing the definition of a problem or opportunity through observation; 2) brainstorming, a formal framework for ideation; 3) prototyping, building models as a source of ideation and the testing of ideas. DT has been linked also to different theories; for example the resource-based view (Borja et al., 2009; Rosensweig, 2011), organizational learning (Beckman and Barry, 2007) and practice-theory (Kimbell, 2012).

One of the major differences between DT in the managerial and designerly discourses is the role of the professional designer. In DT, Brown (2008) and Martin (2009) disconnect design from the designer professions, and Brown (2008) refers to ‘design
thinkers’ whose professional background can vary, stating that people outside of professional design can also have a natural aptitude for design thinking. More recently, the use of DT has been proposed as a way for individuals to release their ‘creative confidence’ (Kelley and Kelley, 2013). On the other hand, it has been suggested also that professional designers should play a central role in using and spreading DT, since they have a natural ability for DT, and could take a more strategic role in the organization (Brown, 2009; Liedtka and Ogilvie, 2011). Brown (2009) claims that the outsourcing of the design function to external agencies, which has become common practice, makes it more difficult for firms to engage in DT because they will have fewer professional designers in-house.

**Representations of design thinking**

Descriptions of DT vary but mostly refer to user-centeredness and a focus on extensive user research in the early stages of projects to gain a thorough contextual understanding of user needs; iterative working, prototyping, a fun mind-set, and learning from failure, etc. (e.g. d.school Stanford, 2013). More detailed descriptions depend on how DT is perceived as a concept (Hassi and Laakso, 2011).

A prescriptive process:

The most tangible representations of DT are linked to IDEO (e.g. Kelley and Littman, 2001; Brown, 2008; 2009; IDEO, 2009), as well as the d.Schools; academic institutions offering DT education for masters students and executives (Stanford d.School, 2013). These organizations propose DT as a process involving a multi-disciplinary team applying a set of design practices to an innovation challenge - an approach that became widespread after ABC Nightline featured the video “The Deep Dive” in 19998. Following these descriptions, DT can be seen as an innovation process consisting of a number of steps (e.g. Kelley, 2001; Stanford d.School, 2009) or a set of “overlapping innovation spaces” (Brown, 2008; 2009; Brown and Wyatt, 2009). Despite some differences in its representation, a generic DT process typically consists of the following steps (figure 1): *Understand* the prerequisites of the problem (the market, the client, technology, perceived constraints); *Observe* users in real life situations using a variety of ethnography techniques to develop empathy for the users; *Define* insights (create a point of view for reframing the problem); *Ideate and prototype* multiple alternatives in short iterations; *Test* by getting feedback, then modify and reiterate solutions, and if necessary, also problem formulation (Kelley and Littman, 2001; Brown, 2009; Brown and Wyatt, 2009; IDEO, 2009; Stanford d.school, 2009).
It has been stressed that DT should not be considered a linear process since a project can move back and forth between different phases, and since ideation, creation of prototypes, and testing and adapting prototypes are described as highly intertwined activities (e.g. Brown, 2008; 2009; Liedtka and Ogilvie, 2011). Best practice includes a dedicated space for creativity and visualization (Brown, 2009; IDEO, 2009; Stanford d.school, 2013).

![Figure 3.8: Description of a DT process (source Stanford d.School, 2009).](image)

Design methods and practice:

Common visualization and prototyping methods described in relation to DT include techniques such as sketching, building scrap models, acting, role-play, storyboarding, storytelling, personas, metaphors and analogies (Stanford d.school, 2013; Liedtka and Ogilvie, 2011). Using the walls of a project room, or a “creative space” to make sense of large amounts of data is described as common practice (ibid). A recurring theme in the DT literature is co-creation with users. In the context of DT, practices such as iterating concepts and unfinished prototypes with users in short loops are described, and involving users actively in an empathy-building phase, for example by inviting them to communicate visually in various ways (Brown, 2008; McCreary, 2010; Lin et al., 2011; Liedtka and Ogilvie, 2011; Stanford d.school, 2013).

A specific mindset:

Linked to ideas of DT as a cognitive matter (Martin, 2009; Boland and Collopy, 2004), a specific DT mindset is central to descriptions of DT linked to IDEO and Stanford. Brown (2008, p.87) describes a design thinker as someone who has empathy – and “can imagine the world from multiple perspectives – those of colleagues,
Innovation

clients, end users, and customers”. He argues “great design thinkers observe the world in minute detail. They notice things that others do not and use their insights to inspire innovation”. Brown argues further that design thinkers are characterized by integrative thinking – not only relying on analytical processes “those that produce either/or choices” – and are able to see all aspects of a problem. Design thinkers are said to be optimistic – any constraints are seen as positive, leading to better solutions – and experimental by posing questions and exploring constraints “in creative ways that proceed in entirely new directions” (Brown, 2008, p.87). Finally, design thinkers are described as enthusiastic about collaboration with individuals from other disciplines and interdisciplinarity.

Areas of application:

DT is often being described as a generic approach to problem solving that can be applied to any situation that any organization or community might face such as formulating business strategy (Holloway, 2009), organizational renewal (Sato et al., 2010) and other “areas of life” (Stanford d.School, 2013). It has been described as a culture or set of principles to guide employees in everyday work (Martin, 2011). When DT is described in relation to NPD or innovation work, it often focuses on the front end of innovation (e.g. Martin, 2009; Lockwood, 2009; Stanford.dSchool, 2013). There are other descriptions that include ‘implementation’ or ‘delivery’ as the final stage in an ‘action plan’ (e.g. Kelley and Littman, 2001; Brown, 2009; IDEO, 2009) or a learning launch (Liedtka and Ogilvie, 2011) signaling its use also in later stages of innovation work. However it is not always clear what is implied by such an action plan or learning launch. Innovation scholars, who study use of design methods or DT in an innovation context, often examine the front end of innovation (e.g. Seidel and Fixson, 2013; Bessant and Maher, 2009).

Overlap with related fields

Similar to any emerging management concept, the ideas held forward are seldom completely new. In isolation, many of these processes and mindsets have been discussed in other research fields, for example innovation and NPD. Some examples are creativity in multidisciplinary teams (e.g. West 2002), ethnographic user research (Leonard and Rayport, 1997), co-creation with users (e.g. Von Hippel, 2009; Jansen and Dankbaar, 2008; Öberg, 2010), market learning in terms of prototyping and user feedback (e.g. O’Connor et al., 2008), and use of analogies (Kalogerakis et al., 2010).

Brown (2009) refers to prototyping as ‘building to think’, and according to Liedtka and Ogilvie (2011, p.49), “visualization is the mother of all design tools” – not just
for visualizing concepts, but also for making an idea tangible. Hargadon and Sutton (1997) as well as Mascitelli (2000) find that drawings, models, and prototypes are useful to product developers as they evaluate and refine new ideas in early phases of innovation work. They note that in early work, it is less about validation of ideas and more about sparking imagination or facilitating understanding between individuals from different functions or professional backgrounds.

According to Mascitelli (2000), breakthrough innovation results from making use of the tacit knowledge and creative energy of individuals/project teams. The goal should be to “establish a generative atmosphere for breakthrough innovation, in which divergent thinking, improvisation, and artistic creativity merge with the practical demands of the product development process.” (ibid:179). Hobday et al. (2012) find a resemblance between descriptions of DT and previously proposed practices and concepts; for example Lindblom’s (1959) “science of muddling through” as an alternative to a rational approach under conditions of uncertainty and incomplete information.

**Critique of design thinking implementation in different settings**

The concept of DT has some critics in the design research community (e.g. Tonkin-wise, 2011; Johansson-Sköldberg et al., 2013; Jahnke, 2013). First, the managerial discourse on DT has been accused of presenting the concept as something that will create value in any setting, and is straightforward to implement. The ease of implementation is contradicted by studies on the integration of design in for example NPD and marketing (Persson, 2005; Perks et al., 2005; Persson et al., 2007), which is often linked to a clash of logics (analytical, rational vs. interpretative, intuitive), as argued by e.g. Rylander (2009) and Edelholm (2007).

According to Johansson-Sköldberg et al. (2013) there are some dimensions that are missing when designerly thinking is translated into DT. First, they argue, design methods are often taken out of context, and do not consider the “embodied knowledge” that is important to designers. Presenting various design tools as a toolbox from which one can pick and choose, regardless of skill, leaves out the knowledge needed to use these tools with competence, which, according to Johansson-Sköldberg et al. (2013), requires years of training, and is embodied in designers. Kimbell (2012) also criticizes managerial discussion on DT for claiming to take inspiration from how designers think and work, generalizing the competences of all designers.
4 Research process

Introduction

In this chapter, I evaluate relevant research approaches and describe the methodology used to collect data for investigation. The objective of this chapter is to provide assurance that adequate procedures were followed. The chapter is organised as follows: first, the philosophical assumptions of positivism and interpretivism are evaluated. Next, research approaches are revised and Gathers’ (1992) taxonomy of research approaches in information systems, is used to provide guidance on selecting appropriate research methods. The main purpose of the framework is to provide guidance to researchers on selecting an appropriate and well-justified research strategy, which is consistent with the aims and objectives of the study.

Third, the selected research method for this study is presented and the limitations of the research method approach are discussed. Next, the research design and the data sources are described in order to explain how they contributed to the achievement of the research objectives. Finally, the steps used in conducting the empirical work are explained, and the grounded theory procedures for analysing data are reviewed.

4.1 Philosophical assumptions of positivism and interpretivism

This section provides an overview of the underlying theoretical assumptions of the two core research philosophies that have been used to conduct research in innovation management. These two methods are extensively reviewed in literature and hence the discussion here will be brief with the aim of gaining insights into the different philosophies and assumptions that could be used for conducting this research.

When the core of the research shifts away from technological to managerial and organisational issues, interpretive research is more popular (Myers, 1997; Hirschheim, 1992; Benbasat et al., 1987) and there is an increasing desire to study problems in
the richness of their real-life setting as opposed to the artificial context of laboratory settings. Interpretive in-depth case studies are prominent among the variety of interpretive research strategies that have been suggested in the research literature, and include (Walsham, 1993; 1995), grounded theory (Glaser and Strauss, 1967), ethnography (Harvey and Myers, 1995), phenomenology (Boland, 1985), hermeneutics (Lee, 1994), and critical hermeneutics (Myers, 1994; 1995).

The aim of both positivist and interpretivist perspectives are to provide a greater understanding of given phenomena, but the approaches are dramatically different in regards to the way this is achieved. The positivist and interpretivist approaches are therefore required to not be viewed as irreconcilable and can be regarded as two approaches that provide different views of the same phenomenon. In addition, Positivist and interpretivist methods can be present simultaneously and mutually supportive within one study (Kaplan and Duchon, 1988; Lawrence, 2002).

4.1.1 Positivist perspective

IS research can be classified as positivist if there is evidence of formal propositions, quantifiable measures of variables, hypothesis testing, and the drawing of inferences about a phenomenon from a representative sample to a stated population (Orlikowski and Baroudi, 1991). The approach is on the basis of 'rational' positivist thought that comprises of the presumption that there is a 'real world', which exists independently of human existence or their perception of it, observation of the phenomena under investigation can be conducted objectively and rigorously, events in this world are determined by natural laws, and that insights arise from discovering these natural laws. Positivist researchers are inclined to not take into consideration the fact that people think and act, and that people are active makers of their physical and social reality. The primary data collection techniques for this perspective are sample surveys and controlled experiments, and inferential statistics is the data analysis method used to 'discover' causal laws. The validity and reliability of identifying and measuring instruments, researcher detachment from the research process, random assignment of subjects and control over confounding influences, are critical to this approach (Orlikowski and Baroudi, 1991; Lawrence, 2002).

4.1.2 Interpretive perspective

A major difference between the interpretive and positivist worldview is the interpretive primary presumption of social constructivism. The underlying premise of the interpretive researcher is that "individuals act towards things on the basis of the meanings that things have for them, that meanings arise out of social interaction, and
that meaning are developed and modified through an interpretive process" (Boland, 1979; Orlikowski and Baroudi, 1991). The interpretive philosophy, emphasizes that the positivist assumptions are unjustified, and that facts and truth are 'chimera', objective observation is impossible, and the act of observation and interpretation is dependent on the perspective adopted by the observer (Clarke, 1994).

Interpretive approaches do not have the assumption that there is an objective reality that can be discovered and replicated by others, in contrast, the approach is based on an ontology in which reality is subjective, a social product constructed and interpreted by humans as social actors according to their beliefs and value systems (Darke et al., 1998). Interpretivism states that reality, in addition to our knowledge of it, are social products and consequently incapable of being understood independent of the social actors (including researchers) that construct and make sense of that reality. The world is not conceived of as a permanent and unchanging constitution of objects, but rather as "an emergent social process as an extension of human consciousness and subjective experience" (Burrell and Morgan, 1979).

In interpretivist research the notion of value-free research is rejected and it is not concerned with the repeatability of an explanation. The value of explanation is judged in terms of the extent to which it allows others to understand the phenomena (Walsham, 1995). The approach is typically concerned with in-depth understanding of the 'real world' issues through accessing the meanings that participants assign to them and focuses on their cultural and historical context (Orlikowski and Baroudi, 1991). The approach contends that theories and concepts are typically risen from the enquiry (Robson, 1993) and it does not redefine dependent and independent variables, and rather focuses on the complexity of humans’ sense making as the situation emerges (Kaplan and Maxwell, 1994; Lawrence, 2002).

Interpretivists argue that organisations are not static and the relationships between people, organisations and technology are continuously changing and are not fixed. The interpretive philosophy, assumes that the social world is not 'given', but it is produced and reinforced by humans through their actions and interactions, and that organisations, groups and social systems do not exist separately from humans, and hence cannot be apprehended, characterised, and measured in an objective or universal way (Orlikowski and Baroudi, 1991). The interpretivist researcher aims to obtain an in-depth understanding of the phenomena being investigated, and acknowledges his or her own subjectivity as part of the process. It allows participants to use their own words and images, and to draw on their own concepts and experiences.
Based upon the interpretive assumptions and given that this study is concerned with the successful use of data analytics for driving innovation in the complex social context of an organisation, the general characteristics of interpretivism become evident as the most appropriate perspective to use in this thesis. I believe that the best way to understand a complex phenomenon such as data analytics enabled innovations, is to adopt an approach that is concerned with understanding the experiences of the participants involved, and therefore the interpretive viewpoint is selected. The primary aim of this study is to describe, interpret, analyse, and understand the factors required for the successful use of data analytics to derive innovation in large organisations. Therefore, given that the creation of innovation and value in large organisations is a dynamic and social phenomenon, it requires an appropriate approach to assist me in understanding the complex interactions within their natural settings.

The interpretive perspective is intended to aid researchers in understanding people and the social and cultural contexts within which they live. According to Klein and Myers (1999) the interpretive approach can take advantage of research methods such as ethnography, case study and grounded theory, that provide explicit recognition to the world of consciousness and humanly created meanings. In this thesis, I agree with Burrell and Morgan (1979) that reality is socially constructed, thus I avoid imposing externally defined categories on this study and all the categories emerged and were grounded in the data (Strauss and Corbin, 1998). Therefore, the categories and themes that emerged in this study were closely coupled to those relevant to the study’s participants. Moreover, I agree with the criticisms against scientific approaches with respect to the objectivity of observations and impartiality of the researcher. Nonetheless, the positivist approaches are still invaluable in generating data about situations and the difference is in the analysis of the data. I used the grounded theory approach for interpreting the situation and building theory, as opposed to relying on quantitative analysis.

4.2 Research approaches

A selected research design must be suitable for the phenomenon under investigation. Thus the nature of innovation and the factors that influence the success of data analytics enabled innovation will have implications for the choice of the suitable methodology. In the next sections I will review a number of research approaches.
4.2.1 Quantitative research

Quantitative research designs are characterised by the assumption that human behaviour can be explained by what may be termed "social facts", which can be investigated by methodologies that use "the deductive logic of the natural sciences" (Jones, 1997). Quantitative research investigates the "distinguishing characteristics, elemental properties and empirical boundaries" and typically measures "how much", or "how often" (Nau, 1995; Lawrence, 2002).

A quantitative research design is flexible in the treatment of data, in regards to comparative analyses, statistical analyses, and repeatability of data collection, in an attempt to verify reliability. Even though this approach is undoubtedly useful in determining the extent of behaviours or attitudes, the adopted methodology fails to provide any explanation or analysis beyond the descriptive level.

Jayaratne (1993) proposes some advantages of a quantitative research design, suggesting that in addition to producing what may be considered as more objective data, it may also allow more objective analysis. Therefore, quantitative methodologies have strengths for IS research and these advantages may be summarised as follows: appropriate for measuring overt behaviour; strong in measuring descriptive aspects; allow comparison and replication; more objective determination of reliability and validity than qualitative techniques.

These strengths however, are not the sole prerogative of quantitative designs. Undeniably, many of the arguments for the use of quantitative research, particularly in an academic climate where resources are limited, have practical origins in terms of allowing large scale data collection simply and inexpensively in a variety of settings, and analysis at reasonable cost and effort, in addition to providing statistical "proof. The weaknesses of quantitative research designs however, is that they mainly fail to ascertain deeper underlying meanings and explanations of the phenomenon under investigation.

The quantitative assumption is that the elements resulting in innovation through data analytics can be reduced to a set of variables, that are equivalent across organisations and across situations. Quantitative research is strong in measuring such variables and, if the focus of the research was the measurement of those variables, then a quantitative approach could have been justified. However, attitudes are important to creation of innovation in organisations and although quantitative methods can be used to measure such factors, they are considerably less appropriate in providing an in-depth explanation of them. Furthermore, quantitative approaches typically take a "snapshot" of a situation - they measure variables at a specific moment in
time- which further adds to the weakness of this method in providing an in-depth explanation of a given phenomenon.

### 4.2.2 Qualitative research

Qualitative research designs are associated with interpretative approaches rather than measuring discrete, observable behaviour and are used to answer questions about the nature of phenomena with the purpose of describing and understanding the phenomena from the informant’s point-of-view (Leedy, 1997). Qualitative methodologies possess strength in areas that have been identified as potential weaknesses within the quantitative approach. For example interviews and observations can be used to provide and obtain a deep, rather than broad, knowledge and understanding about a particular phenomenon, and the attitudes of the organisations towards enabling innovation through data analytics can appropriately be investigated using qualitative research.

The argument used in this thesis is that quantitative methods measure human behaviour "from outside", without accessing the meanings that individuals give to their measurable behaviour. Conversely, a qualitative research design permits understanding to be gained from the informants’ point of view. The advantages of a qualitative methodology for the phenomenon under investigation are that first they allow the phenomenon to be explored in greater depth than quantitative methodologies; and second they encourage the informant to introduce concepts of importance rather than adhering to subject areas that have been pre-determined by the researcher. Research into data analytics as an enabler of innovation is not extensive and thus the flexibility of qualitative methodologies is appropriate for this research that is exploratory in nature.

Objections to the approach do exist. However, the main argument against it, is the concept of validity, in that it is difficult to determine the truthfulness of findings. The relatively low sample numbers often encountered may also lead to claims of findings being unrepresentative of the population. Furthermore, the selection of cases may lead to the criticisms of the cases being untypical.

### 4.2.3 Taxanomy of research approaches

Galliers (1992) suggests a framework that can be utilised to identify suitable research methods in information systems. Galliers emphasises that the framework should not be used as a prescriptive means for identifying the ‘correct’ method, and rather it should be used only as a guide with the purpose of presenting a starting point for the selection and rejection of methods. Gallier’s framework, which is presented in
Research process

Table 4.1, is used in this study as a guideline for selecting an appropriate research methodology.
<table>
<thead>
<tr>
<th>Object</th>
<th>Mode for traditional empirical approaches (observations)</th>
<th>Modes for newer approaches (interpretations)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theorem Laboratory Proof (experiment) Field</td>
<td>Forecasting and futures research Simulation and game/ role playing Subjective/ argumentative (including reviews) Descriptive interpretive Action research</td>
</tr>
<tr>
<td></td>
<td>experiment experiment Case study Survey</td>
<td>research playing</td>
</tr>
<tr>
<td>Society</td>
<td>No No Possibly Possibly Yes Yes Yes Yes Possibly</td>
<td></td>
</tr>
<tr>
<td>Organisation/ group</td>
<td>No Possibly (small groups) Yes Yes Yes Yes Yes Yes Yes</td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>No Yes Yes Possibly Possibly Yes Yes Yes Yes Possibly</td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td>Yes Yes Yes No Possibly Yes Yes Possibly Possibly No</td>
<td></td>
</tr>
<tr>
<td>Methodology</td>
<td>Yes No Yes Yes Yes No Yes Yes Yes Yes Yes Yes</td>
<td></td>
</tr>
<tr>
<td>Theory building</td>
<td>No No No Yes Yes Yes Yes Yes Yes Yes Yes Yes</td>
<td></td>
</tr>
<tr>
<td>Theory testing</td>
<td>Yes Yes Yes Possibly No Possibly No Possibly Yes Yes</td>
<td></td>
</tr>
<tr>
<td>Theory extension</td>
<td>Possibly Possibly Possibly Possibly Possibly No No No Possibly Yes</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Information systems research approaches (Galliers, 1992)
As depicted in Table 4.1., Galliers categorises the information systems research approaches into two main categories of scientific and interpretive. The scientific category includes traditional empirical approaches that are based on ‘rational’ positivist thought that assumes observations of phenomena under investigation can be made objectively and rigorously. The interpretive category on the other hand, contains newer approaches that are based on subjective interpretations. Interpretivist argues that the scientific philosophy is misplaced in social scientific enquiry due to several reasons: 1) there is the possibility of various interpretations of social phenomena; 2) the impact of the social scientist on the social systems being studied; 3) the complications associated with forecasting future events concerned with human activity, considering that there will continuously be a mixture of intended and unintended effects and 4) the risk of self-fulfilling predictions or the opposite (Galliers, 1992).

The taxonomy reviews the probable appropriateness of each method in the context of the research topic under investigation. It indicates several potential research methodologies that can be used in conducting research in information systems, including laboratory and field experiments, forecasting/futures research and simulation and game/role playing, surveys, case studies, action research and observation. Below a brief discussions of several of these research methodologies that are described in the framework, is provided.

**Laboratory and field experiments** that assume real-world interference is not significant to the events being studied, permit the researcher to isolate and control a small selection of variables, and study them intensively. Because the factors that influence the use of data analytics for innovation are interpreted in complex situations, it may not be valid to attempt (1) to isolate change within an experimental situation due to the fact that this assumes a simplified snapshot view with a well-defined beginning and an ending (Pettigrew, 1985; Vitalari, 1985). (2) To develop an explanation based on models of abstracted causality (Galliers and Land, 1987). Furthermore, since the context of each situation is different for each case, I am unable to predict precisely how the participants will interpret the factors that influence their success in the use of data analytics for innovation. This makes it infeasible to conduct controlled experiments since I cannot know in advance which variables to control, manipulate or to exclude (Baskerville and Wood-Harper, 1992). Consequently, laboratory and field experiments are believed to be unsuitable for this study as the variables regarding the success of data analytics for innovation, are not presupposed and identified.

**Forecasting/ future research, and simulation and game role playing**
Forecasting/future research has the potential of delivering insights into future events, however it has limitations such as self-fulfilling predications and the unpredictability of environmental factors for epistemological reasons similar to those applying to experimental approaches. I believe that the forecasting/future research method is less suitable because this study is not concerned with future uses of the data analytics for innovation; its focus is on how data analytics is being used now in business to derive innovation and how it can also be implemented in other organisations.

Two other potential research approaches are simulations and game/role playing that have the benefit of studying situations that might otherwise be impossible to analyse. These methods have an additional limitation compared to forecasting/future research, where they are not based on real-life interventions. Therefore, these approaches are unsuitable for use in this thesis as the objective of this study is not to extrapolate events, and is instead to describe and understand the necessary constructs for the use of data analytics for innovation now.

**Action research**

This approach is based on the assumption that the researcher cannot be disconnected from social debate and therefore it aims to make their biases explicit and record all the successful/unsuccessful interventions (Lyytinen and Klein, 1985). Action research involves not only observing and recording but also actively taking part in endeavours to solve problems on site and essentially intervening in a situation. The first aim of this PhD is to determine the success factors for creating innovative products and services through data analytics, and building a theory for this. For this goal, I believe action research is inappropriate because I am not seeking to intervene in a research context to achieve an outcome. However, for the second part of this PhD research that aims to validate the CDA framework created from the theory, action research is appropriate as it allows me to record actions that result in successful or failed outcomes and I can take an active role in enabling these outcomes. Thus, the second part of this research is concerned with intervening or correcting actions in a project setting; while the first part of this thesis focuses on understanding and making sense of how the social actors use the phenomenon under investigation.

**Observation**

Innovating using data analytics is not a linear process and requires many iteration cycles that means it might take many months if not years, before success is realised. Also, not every project will succeed and experimentation is crucial for innovation
projects. Therefore, pure observation, whether passive or participant, is impractical and unsuitable for this study.

Galliers framework classifies case study in the traditional empirical approach. However, I argue that this method should be classified in both traditional and newer approaches, as its underlying philosophy also fits well with interpretivism. Cavaye (1996) reports that despite the dominance of the positivist style of case study, there is a clear increase in the use of interpretive case study research on IS problems, and papers considering the interpretive case study research approaches have appeared in major IS journals (Walsham, 1995; Klein and Myers, 1999). Cavaye argues that case study is a versatile and pluralistic research method that can be utilised in various ways to suit the objective of the phenomenon under investigation. The case study method can be conducted with a "positivist stance (Yin, 1994; Benbasat et al., 1987) or an interpretivist stance (Walsham, 1993; Myers, 1997), it can take a deductive or an inductive approach, it can use qualitative and quantitative approaches and can investigate one or multiple cases (Cavaye, 1996).

4.3 Research method chosen for this study

4.3.1 Case study method

This study takes the form of exploratory and descriptive research focusing on the use of data analytics for creation of innovative products and services. I conducted a comprehensive examination of Gathers (1992) framework by considering the research methods and their purpose, strengths, weaknesses and their ability to meet the objectives of this study. Given the novelty and changeable nature of the area of study, the research questions were best pursued in the context of a qualitative study.

For this research a multiple-case study was carried out to obtain qualitative information and grounded theory technique (Strauss and Corbin, 1990) was used to analyse the case data. This was essential in order to provide a comprehensive understanding of the issues under investigation. Robson (1993) suggests that when the purpose of the research is exploratory – which is the case in this research - then case study method will be the most appropriate. The case studies are fundamentally different from surveys (and laboratory experiments and field studies) because the researcher typically has less presumptive knowledge of what the variables of interest will be and how they will be measured (Gable, 1994). In this PhD research, the purpose of the case study was to provide an understanding of the factors that influence the success of data analytics driven innovations in large organisations. Although the case study
Research process

methods are classified under the traditional empirical side of Galliers framework, their adoption here is based on epistemological assumptions of interpretivism.

Yin (1984) argues that case studies are appropriate in situations where the goal is to study contemporary events, and where controlling behavioural events or variables is not required. Yin (1994) defines case studies as “an empirical inquiry that investigates a contemporary phenomenon within its real life context, particularly when the boundaries between a phenomenon and context are not clearly evident; and in which multiple sources of evidence are used”. The contemporary phenomenon being investigated in this thesis is the understanding of how data analytics can be used in large organisations to result in innovation. I believe that this phenomenon should be investigated in its real-life context, since all the variables in this context contribute to its understanding.

Furthermore, this methodology is suitable in situations where an understanding of interactions between information technology-related innovations in organisational contexts is required (Darke et al., 1998). Case study can be used in different ways using various research perspectives and data collection and analysis methods; to produce diverse types of research outcomes (Cavaye, 1996). Darke et al. (1998) argue that the case study method is applicable for providing descriptions of phenomena, exploration, prescription, theory building or testing existing theoretical concepts and relationships (Benbasat et al., 1987; Cavaye, 1996).

Orlikowski and Baroudi (1991) claim that case study methods are appropriate in generating valid interpretive knowledge, since these methods examine humans within their social setting. Moreover, the case study methods are especially suitable when theoretical knowledge on the phenomenon under investigation is limited and an understanding is not well established (Benbasat et al., 1987). These include areas where a phenomenon is dynamic and not yet mature or settled, such as the use of data analytics for innovation and its influencing factors, where there are limited existing theories to explain the phenomenon (Hutchinson, 1988).

The case study method provides many advantages that include (Galliers, 1992; Yin, 1989; 1994): the ability to capture realities in greater detail by studying a phenomenon in its natural context; analysing more variables than is feasible using other methods (Benbasat et al., 1987); and asking penetrating questions and capturing the richness of organisational behaviour that is valuable in developing and refining concepts for further study. The case study method also provides rich and explanatory evidence that can be used to explain why and how phenomena occur and it is appropriate for studying topics where attitudes and behaviours can be understood
Research process

best within their natural setting. Similarly, Benbasat et al. (1987) claims that the case study method has three strengths when used in information systems research: (1) Study information in a natural setting, acquire knowledge about the state of the art, and generate theories from practice; (2) appreciate the nature and complexity of the process taking place; (3) gain valuable insights into novel topics emerging from rapidly changing fields such as innovation.

Clearly, the case study method also has various disadvantages. These limitations will be further discussed in the next section.

4.3.2 Limitation of research method

As expected with all research designs, the approach selected in this study has both strengths and weaknesses. It is essential to acknowledge these weaknesses in order to ensure that the study’s findings can be placed in an appropriate context, in addition to ensuring that future research is able to improve upon the present design. Several problems have been recognised in regards to the research methods utilised in this study.

Although the case study method is advantageous as a means of examining the factors that influence the success of data analytics driven innovations in large organisations, there are a number of practical challenges associated with attempting to undertake case studies as rigorous and effective methods of research, regardless of the adopted philosophical perspectives (Darke et al., 1998). In general, it may be difficult to design and scope a case study project to ensure that the research questions are answered appropriately and adequately. Despite the efforts to gain the greatest quality of information possible, there were still several limitations involved, such as: lack of access to specific projects and financial data on these projects since most participants were not prepared to reveal these data; restricted availability of appropriate case study sites since organisations are not always willing to participate in case study research; the amount and variety of data from case study may inhibit data analysis, especially as strategies and techniques for analysing qualitative data are not well established (Miles and Huberman, 1984; 1994; Yin, 1994; Cavaye, 1996).

In case study method, both the data collection and analysis processes are susceptible to the influence of the researcher’s characteristics and background, and rely heavily on the researcher’s interpretation of events, documents and interview material (Galhers, 1992) that may limit the validity of the research findings. These diverse interpretations of events by researchers are typically attributed to the author rather than the approach, because the results of case studies are significantly dependant
on the way they are conducted (Galliers, 1992). Kaplan and Duchon (1988) argue that the understanding of reality is based on the researcher’s interpretation of the data and therefore researchers may interpret the same data in different ways.

Another potential challenge that needs to be considered is the bias introduced by the researcher during the collection and analysis of case data. The interpretive researchers recognise the subjectivity of their analysis in that their predisposition, beliefs, values and interests continuously intervene to shape their investigations (Orlikowski and Baroudi). Darke et al. (1998) describe two types of biases: (1) the effects of the researcher on events and the behaviour of participants at the case study site. (2) The researcher’s own beliefs, values and prior assumptions that may prevent adequate investigation and consideration of potential contradictory data, and improperly influence the analysis of the case study evidence. According to Yin (1994) bias exists in all research methods to varying degrees and may influence the design and conduct of any type of research methods, and therefore researchers must conduct any research with a high level of care.

In this study, in an attempt to counteract my biases in collection and analysis of case data, multiple sources of evidence were used to provide multiple instances from different sources (Miles and Huberman, 1984; 1994). When faced with inconsistencies or illogical arguments during interviews, I investigated for further explanation. I compared the research findings with other similar studies in this area (especially in the field of innovation) for inconsistencies and I also discussed the findings with colleagues and my research supervisor in great detail. Convergence of information from a variety of sources also provided a means of strengthening the convergence of information, giving multiple measures of the same phenomenon (Yin, 1994). This was also helpful for verifying information provided by different participants where there were conflicting accounts of events and actions.

Another criticism of the case study method is lack of generalisation, where case studies are generalisable to theoretical propositions and not to populations (Yin, 1988). Theoretical statements can be developed using case results and can subsequently be used to build theory. Case results can yield data from which process theories and richer explanations of how and why processes and outcomes occur can be developed (Kaplan and Duchon, 1988). In addition to Yin (1988), Walsham (1993) further acknowledges that criticisms of case study method for empirical research are typically focused on the non-representativeness and lack of statistical generalisability that arise from such research. This is occasionally overlooked if the study is used as an exploratory method of analysis prior to, or in addition to, more comprehensive large
sample work (Walsham, 1993). In this study however, this is not the justification for the use of the case study method.

The principal argument for the use of the case study method in this research, as discussed by Smith (1989), is that epistemology is interrelated. If a positivist epistemological stance is adopted, then statistical generalisability is the main objective, however, in this research I have adopted an interpretive epistemological stance. Walsham (1993) reasons that the validity of the inferences obtained from one or more cases does not depend on the representativeness of cases in a statistical sense, "but on the plausibility and cogency of the logical reasoning used in describing the results from the cases, and in drawing conclusions from them". Walsham argues that there are four types of generalisations from interpretive case studies: the development of concepts, the generation of theory, the drawing of specific implications, and the contribution of rich insight (Walsham, 1995).

In this research, generalisation from the setting to a population is not pursued, rather, the intent is to understand the factors that influence the success of data analytics enabled innovations in large organisations, in order to create a theoretical model of that process. This study is attempting to create a theory that is compatible with evidence that is both rigorous and relevant and largely beneficial to other areas. This is more in line with the constructivist criteria for quality of research, which relies upon the richness or authenticity of the learning that is achieved (Hughes and Wood-Harper, 1999).

4.4 Data collection

This section provides an explanation of how the multiple case studies were designed, and how the data were collected. The aim at the data gathering stage is to generate the richest possible description of the phenomena being studied, because regardless of how efficient the subsequent analysis is, thin data will result in thin outcomes that are not likely to reflect the substantive issues involved. Kling (1991). Danziger and Kraemer (1991) argue that to adequately address the impact of information technology, it is best to exploit multiple methods of data collection. Using several techniques can result in novel insights and modes of analysis that are not likely to occur if only one technique is utilised (Kaplan and Duchon, 1988; Lawrence, 2002). In this study, data was collected through various methods that included semi-structured open ended interviews and other secondary documents. A further source for comparison was the reviewed literature.
This triangulation across different techniques of data collection was especially beneficial in theory generation, because it provided multiple perspectives on each issue, supplied more information on emerging concepts, allowed for cross-checking, and produced stronger substantiation of constructs (Eisenhardt, 1989; Glaser and Strauss, 1967; Pettigrew, 1990; Orlikowski, 1993). Data collection, coding, and analysis proceeded iteratively (Glaser and Strauss, 1967), with the initial stages of the research being more open-ended, and subsequent stages being directed by the emerging concepts. As noted by Eisenhardt (1989), this overlapping of data analysis and data collection, provided a number of advantages: "it not only gives the researcher a head start in analysis, but more importantly allowed the researcher to take advantage of flexible data collection. Indeed, a key feature of theory-building case research is the freedom to make adjustments during the data collection process" (Orlikowski, 1993).

4.4.1 Case study design

In this study I used an emergent, exploratory, inductive qualitative case study approach because it does not predetermine or define the directions the investigation might take. According to Lincoln and Guba (1985), the emergent design of a naturalistic inquiry does not permit a detailed plan to be created prior to the start of research "the research design must therefore be 'played by ear'; it must unfold, cascade, roll and emerge". The design of the case study in this research, commenced with a broad outline of contingency plans that were open to modification and extension as required during the study (Cantrell, 1997). The design assumed a worldview in which there are multiple realities that are complex, dynamic and change overtime, and that the world is not objective but rather a function of interactions and perception (Merriam, 1988). This view is central to this research that aims to bring understanding and interpretations to processes or events as perceived by organisations in their natural setting.

To investigate how large organisations create innovative products and services through the use of data analytics and the success factors for such innovation, I conducted a multiple case study with 12 of the most innovative and largest organisations in the world. The case studies were conducted using semi-structured face to face interviews in the organisations’ offices in Silicon Valley.

As previously mentioned, case study method involves conducting research into the natural setting of the phenomena under investigation. According to Yin (1984), case study can include single or multiple cases and many levels of analysis. He further argues that multiple case designs are desirable when the intent of the research is
descriptive, theory building or theory testing. Furthermore, Benbasat et al (1987) state that such an approach is suitable for investigating "certain types of problems: those in which research and theory are at their early formative stages; and sticky, practice- based problems, where the experiences of the actors are important and the context of action is critical". They claim that multiple case studies enable the researcher to relate differences in context to constants in process and outcome, and in addition, multiple cases allow for cross case analysis and the extension of theory. Miles and Ruberman (1984) further note that multiple cases enable the researcher to confirm that findings are not merely the result of the characteristics of the research setting.

4.4.2 Case participants

The first innovative company was selected at random from the organisations that participated in this study, to provide the first body of data. Subsequent data collection was guided by the theoretical sampling principle of grounded theory as defined by Strauss and Corbin (1990). Therefore sampling was conducted on the basis of concepts that proved theoretical relevance to the evolving theory because in a grounded theory theoretical sampling cannot be fully planned before the study commences.

Request for an hour long interview with explanation of the nature of the research was provided to the participating organisations. List of all participating organisations is provided in Chapter 5 in Table 5.1.

Interview

Semi-structured open ended interview was the principle tool for collecting data for the case study. Interviews are debatably the main data sources where interpretive case study research is undertaken (Yin, 1994), because interviews enable researchers to gain great access to case participant’s views and interpretations of actions and events (Walsham, 1995). According to Kaplan and Maxwell (1994), the main purpose of an interview is to obtain the respondent’s views and experiences in his or her own words, instead of collecting data that are merely a choice among pre-established response categories. Interviews possess the adequate flexibility to adapt to each context and organisation, and to pursue unexpected paths and prompts suggested by the theoretical sensitivity (Glaser and Strauss, 1967) that is developed by the researcher throughout the research process.

I selected interview as an appropriate form of data collection in the case study, because the technique enabled face-to-face contact with the social actors and provided
the additional advantage of allowing me to recognise and process non-verbal communications. The interviews were structured to gather data about the widest possible range of issues associated with the phenomenon under study i.e. innovation through data analytics. The research questions guided the data-gathering process and the structure and content of subsequent interviews was determined after the data analysis process had commenced. The interviews were used to (1) gather new data about known concepts and categories that have been developed about the phenomenon, and (2) involve the organisations in a process of testing and verifying data and the emerging theory. The use of interview enabled me to investigate and explore particular responses and also develop insights into how the participants interpret and make meaning of the world.

Interviews were held with managers of data analytics or innovation departments and each interview began with a concise discussion about the aim of the research. The interviewee was encouraged to respond by a number of broad questions, however they were stimulated to offer their own view of the phenomena and to range more broadly than would be the case in a typical structured interview. These questions focused on gaining an understanding of the factors that influence the success of data analytics efforts resulting in innovative products and services.

I used an interview schedule guide that started by asking questions of the ‘grand tour type’ (Spradley, 1979) in order to uncover the background details of the participants and also to relax them by easing them into the interview situation. The interview guide is presented in Appendix 2 and included multiple open-ended questions to allow the participants a superior degree of freedom to offer their answers in their own terms. In these open-ended responses the emerging concerns and issues were discovered, which were compared among respondents, and concepts that evolved were the basis for further data gathering, while consistently leaving room for other answers and concepts to emerge. In addition, I encouraged open discussions toward the end of each interview in order to allow the interviewees to ask questions and add any comments they might desire. Coding was conducted during the gathering and processing of interview notes, tapes and transcripts.

The length of each interview was approximately one hour and they were tape-recorded and subsequently transcribed prior to the commencement of data analysis. Tape-recording of interviews is frequently recommended (Miles and Huberman, 1994; Yin, 1994; Darke et al., 1998) as a means of providing a complete description of the interviewees’ response and comments. However, tape-recording – similar to any other method- also poses some disadvantages such as inhibiting the interviewee
and also the reliance on tape recordings can prevent the researcher from listening carefully and participating fully in the interview process (Darke et al., 1998).

To improve the validity of the responses, abstracts of the main findings of each interview were verified with participants and once a confirmation was received, a matrix was created to allow within and across case analyses. This form of data analysis enabled the researcher to seek out and verify themes within text. As all of the interviewed companies are widely recognised and innovative companies such as Facebook and Apple, a large body of knowledge and documents exist about their innovation strategies and specific innovation case studies where they successfully used data analytics for innovation. Documentations supporting interviewees’ success in achieving data analytics enabled innovations, were obtained and the interviews were supplemented by these secondary data sources such as company documents, journal and newspaper articles. Themes covered during the Interviews included background details of how the firm supports innovation, their data analytics capabilities and their use, success factors influencing the creation of innovative products and services using data analytics, and barriers to success.

The format and approach of the interview were tested and refined in the pilot study where the data-gathering methods were fine-tuned and the theoretical sensitivity towards the phenomenon was heightened. According to Strauss and Corbin (1990) “Theoretical sensitivity refers to the attribute of having insight, the ability to give meaning to data, the capacity to understand, and capability to separate the pertinent from that which isn’t”.

Secondary data

In addition to the primary data collection through semi-structured open ended interviews, I also collected secondary materials as appropriate for each case. These documents included a brief history of the organisation, reports, newsletters, promotional material, Internet web sites and other internal publications. These materials delivered a wealth of information -some of which were not available through interviews - that helped in confirming information from other sources and provided different perspectives on similar information.

In fact, document analysis and texts provided valuable sources of qualitative data (Miles and Huberman, 1994; Kaplan and Maxwell, 1994) by offering additional information as well as verifying and corroborating the information gained through interviews (Lawrence, 2002).
Secondary data was also obtained from literature review of innovation success factors in large organisations.

4.5 Data analysis

This section discusses data analysis in qualitative case study research. One of the principal challenges in conducting qualitative research is selecting an appropriate framework within the boundaries of which the data can be collected and analysed and it can also be used as a starting point. In quantitative research, it is often feasible to make a well-defined distinction between gathering data and analysing data, but, such clear distinction is not applicable to qualitative research. As Tesch (1990) states, in qualitative research, analysis of data and collection of data are conducted simultaneously and they inform and/or drive each other.

Conducting qualitative analysis for the first time presents challenges in regards to both the number of methods and the difficulty of analysing and presenting large amounts of data. Qualitative studies typically generate large amounts of data – often in non-standard formats- that are not easily responsive to automated manipulation, analysis and data reduction (Yin, 1984). Qualitative analysis allows researchers to obtain information and gain insights that may have been neglected and un-noticed with traditional data analysis techniques. In this research the analysis of the multiple case studies was conducted according to the guidance provided by many scholars in this field (such as Glaser, 1978; Glaser and Strauss, 1967; Lofland and Lofland, 1984; Taylor and Bogdan, 1984).

In qualitative research, the data analysis process "involves working with data, organising it, breaking it down, synthesising it, searching for patterns, discovering what is important and what is to be learned, and deciding what you will tell others" (Bogdan and Biklen, 1982). Spradley (1979) refers to this analysis as a "systematic examination of something to determine its parts, the relationship among parts, and their relationship to the whole". Miles and Huberman (1984) on the other hand, describe data analysis for qualitative research as a process comprising of three simultaneous activities. According to them, data reduction is the process of selecting, simplifying, abstracting and transforming the new case data and that data collection and data analysis should overlap. This is to provide the researcher with flexibility in data collection procedures, in order to remain open to new ideas or patterns that may emerge (Lawrence, 2002).

As previously stated, there are many techniques available for the analysis of qualitative data and in order to select the most appropriate method for this study, I
researched various approaches such as hermeneutics, content analysis and semiotics (Myers, 1997). These approaches originate from diverse fields and offer the opportunity for gaining different insights on the data. I evaluated these approaches from two perspectives: 1) whether the approach was utilising all features of case study; and 2) whether the philosophy of the approach imposed any pre-existing theories. As Eisenhardt (1989) advises, the theory building research must commence with fundamentally no theory under consideration and no hypotheses to test. This is because preordained theoretical perspectives may bias and restrict the findings.

The selection of grounded theory for this study (Glaser and Strauss, 1967; Glaser, 1978; Strauss, 1987; Strauss and Corbin, 1990; 1998) from a multitude of other qualitative analysis methods is not arbitrary and is because it has been a leading technique for social research (Hughes, 1999) and its use is increasing in innovation research and the IS field. The objective of this study is the development of a conceptual framework that explains the systematic process of achieving innovation through data analytics in large organisations, which fits well with the philosophical nature of grounded theory. Advocates of the grounded theory believe that an approach which is concerned with the meanings, definitions, and interpretations of those being studied, has superior potential for more accurately representing their world and priorities than methods that commence by preconceiving the world and its meaning (De Biirca and McLoughlin, 1996). Consequently, I did not conduct my research at the field with a clear set of constructs and instruments with which to measure the social reality; and instead I derived the categories from the field by in-depth examination and exposure to the phenomenon. The next section defines grounded theory and introduces its procedure for data analysis.

### 4.6 Grounded theory

This section provides an overview of grounded theory (Glaser and Strauss, 1967; Martin and Turner, 1986; Turner, 1983; Glaser, 1978; Strauss, 1987; Strauss and Corbin, 1990; 1998) and offers further reasoning for why the method was selected as the primary qualitative tool for analysing the case study data. The grounded theory was first proposed by Glaser and Strauss in 1967 and has since been used significantly across various social science disciplines. According to Strauss and Corbin’s (1990) definition, a grounded theory is a theory that is discovered, developed, and provisionally verified through systematic data collection and analysis of data relating to a specific phenomenon (Strauss and Corbin, 1990). The grounded theory is particularly useful when conducting qualitative techniques for the first time, because it offers well-defined procedures. The grounded theory methodology is iterative -
requiring a constant movement between concept and data- and comparative, which means a continuous comparison across types of evidence is required in order to control the conceptual level and scope of the emerging theory.

The aim of grounded theory is obtaining a theory that is closely tied with the evidence, in order to increase the probability that the resultant theory will be consistent with empirical data (Orlikowski, 1993; Eisenhardt, 1989). Therefore as the theory emerges it guides the data collection method, coding rationale, integration of categories, abstracting from the data and construction of theory. Many scholars including Turner (1983), Martin and Turner (1986), Strauss (1987), Charmaz (1983), and Strauss and Corbin (1990), have provided guidance on the use of grounded theory method, which is used in many disciplines including innovation research.

Grounded theory is a general style of conducting analysis that is independent of any specific disciplinary perspectives (Strauss, 1987) and what differentiates it from other qualitative research methods, is its emphasis upon theory development (Strauss and Corbin, 1998). A theory is considered grounded if it emerges from data and produces explanations of relationships and events that reflect the life experiences of those people and processes that the researcher is attempting to understand. In addition, traditional qualitative approaches initially collect the data and then proceed to conduct the analysis, however, grounded theory is different because it uses the emerging theoretical categories to shape the data collection while doing the fieldwork (data collection and analysis proceed simultaneously) (Lawrence, 2002).

Furthermore, a requirement of grounded theory is for the researcher to demonstrate theoretical sensitivity (Glaser and Strauss, 1967; Glaser, 1978) by being grounded in technical literature in addition to having insights from personal and professional experience and from collection and analyses of the data (Strauss and Corbin, 1990). Grounded theory encourages researchers to avoid their thinking being influenced by technical literature and standard ways of thinking about the data (Strauss and Corbin, 1990). The interaction between the emergent theory and technical literature is considered when extending generalisations from the study. This is attained by combining supplementary or conflicting analyses into the theory by including them as categories or conditions, or criticising them in terms of what has emerged (Strauss, 1987). The use of grounded theory analysis is originated on the foundation that the generation of theory at different levels is essential for an in-depth understanding of social phenomena (Glaser and Strauss, 1967; Glaser, 1978). Grounded theory is especially suitable for a case study aimed at exploring the factors that influence the success of data analytics enabled innovations as it is advantageous for understanding
contextual elements (Orlikowski, 1993) that constitute the primary focus of the case studies of this research.

A practical disadvantage of grounded theory is that it is an exceptionally labour intensive method, which requires the investment of significant cognitive effort by the researcher. Despite this, I believe that the grounded theory technique is an appropriate methodology to use, particularly because I am required to analyse large quantities of unstructured and semi-structured qualitative data.

**Reasons for using grounded theory for data analysis**

In this study, grounded theory is selected for analysing the case study data, with the goal of creating a descriptive and explanatory theory of factors that influence the success of the creation of innovative products and services through the use of data analytics. Strauss (1987) stresses the appropriateness of the case study approach when combined with grounded theory and states that it "is an inductive, theory discovery methodology that allows the researcher to develop a theoretical account of the general features of a topic while simultaneously grounding the account in empirical data" (Martin and Turner, 1986; Glaser and Strauss, 1967). This generative approach is especially beneficial in this research because the aim of the study is the discovery of a theory that explains the success factors for innovation using data analytics in large organisations.

Grounded theory provides a means of focusing on qualitative material in detail, to systematically develop theories in regards to the phenomena being studied. Turner (1981) proposes that grounded theory is especially appropriate in dealing with qualitative data that is obtained from participant observation, face-to-face interactions including semi-structured or unstructured interviews, and case-study material or documentary sources (Turner, 1981). Predictably, these specific types of inquiry produce large amounts of data, which accumulate in non-standard and unpredictable formats and the grounded theory approach provides the researcher a strategy for examining and analysing material of this nature. Furthermore, another advantage of using grounded theory is that a documented record of the progress of the analysis is generated and therefore it is feasible to trace the derivation of concepts and models by inspecting the data and memos.

Grounded theory is most advantageous in areas in which minimal research has been conducted and therefore it is particularly useful in the area of our research, which is the use of data analytics for innovation, where there is currently no theory for this phenomenon.
Furthermore, another reason grounded theory is a suitable methodology for this study is that it generates a theory that can be used as a foundation for further investigation of this phenomenon and related issues. In subsequent studies, other qualitative research techniques and/or quantitative methods (or a combination of both), can be utilised to test, verify or extend the qualitative proposals that emerge from this research.

Producing accurate and valuable results is a core principle of grounded theory, and therefore it is important to incorporate the complexities of the organisational context into an understanding of the phenomenon, instead of simplifying or ignoring it (Orlikowski, 1993; Martin and Turner, 1986; Pettigrew, 1990). This method of research necessitates that broader contextual issues, which effect the phenomenon under study, be given suitable acknowledgement and be accounted for in the development of theory (Lawrence, 2002).

Several theoretical methodologies stress the importance of organisational context in influencing technology use in organisations. This belief is also held in this research and the use of a grounded theory approach enables the inclusion and investigation of this crucial organisational element. The grounded theory methodology is an inductive approach that is based on the researcher’s ability to systematically collect, code, categorise and analyse data to facilitate “the generation of theories of process, sequence, and change pertaining to organisations, positions, and social interaction” (Glaser and Strauss, 1967).

Data analytics research has fundamentally been concerned with technical aspects of analysing and coping with data such as data mining, statistical analysis, fact clustering, data visualization, natural language processing, text analytics, and artificial intelligence as opposed to investigating it from the perspective of innovation and the elements and processes necessary for achieving innovation through analytics. The inductive, contextual and processual characteristics of grounded theory fit well with the interpretive orientation of this research that aims to generate a context-based description and explanation of the phenomenon, instead of an objective, static description that is conveyed strictly in terms of causality (Boland, 1979; 1985; Chua, 1986; Orlikowski and Baroudi, 1991). More specifically, in this research the goal is to generate a theory that describes and explains the process of innovating using data analytics in terms of an interaction of contextual conditions, actions, and consequences, instead of explaining variance using independent and dependent variables (Orlikowski, 1993).
In the next section, a description of the appropriate procedures for analysing data in grounded theory is provided.

4.6.1 Grounded theory procedures

The distinguishing quality of grounded theory is that it is a general approach for generating a theory that is grounded in data which is systematically gathered and analysed. A constant interaction between analysis and data collection allows the theory to evolve throughout the research, with data analysis continuously guiding future data collection. This section aims to provide detailed descriptions of the key procedures of data analysis in the grounded theory method.

Strauss and Corbin (1990) classify the level of analysis into three categories of: (a) Presenting the data without interpretation and abstraction where the participants articulate their own story; (b) Creating a "rich and believable descriptive narrative" by means of field notes, interview transcripts and the interpretations of the researcher; and lastly (c) building a theory using great degrees of interpretation and abstraction. In this research, I combined the second and third levels of analysis suggested by Strauss and Corbin, in order to provide rich and comprehensive descriptions, which permit others to make adequate contextual judgements to transfer the case study findings to alternative settings. The purpose here is to express multiple constructions of reality as experienced by large innovative organisations that have already used data analytics for innovation purposes.

Performing data analysis in grounded theory comprises of applying specific procedures that when conducted appropriately will produce a theory which is thorough and well-grounded in the data. As Strauss (1987) stresses, these procedures should be regarded as guidelines as opposed to hard or fixed rules, and therefore researchers are required to study these guidelines, and use and modify them in accordance with the requirements of their research. Moreover, Strauss and Corbin (1998) further advise against rigid adherence to any of these procedures as doing so may obstruct the analytic process and suppress creativity.

Data recording can be regarded as a pre-analytic phase of the grounded theory method and it is considered to be vital to the successful generation of grounded theory (Hutchinson, 1988). The grounded theory methodology encompasses coding the assignment of themes and concepts to a designated unit such as sentences taken from an interview transcript. Categories are then created from combining concepts and links between these categories are identified and verified against the data. Subsequently, selective coding attempts to integrate the categories into a the-
ory that explains the phenomenon being investigated. The process of analysis in grounded theory involves coding data (open, axial and selective), memo writing and theoretical sampling.

4.6.1.1 Coding

Coding is the process of breaking up the data from sources such as field notes and interview transcripts and conceptually grouping it into codes that then form the theory that explains what is happening in the data (Glaser, 1978). Codes can be in the form of a simple category label or a more complex one such as a metaphor (Miles and Huberman, 1994) and pose questions such as 'what does this incident indicate?' (Glaser, 1978). Codes are used by researchers to combine and categorise a series of otherwise discrete events, statements, and observations that are derived from the data (Charmaz, 1983).

Open coding is the process of analytically identifying concepts and discovering their properties and scopes in data, which is conducted by naming and categorisation of phenomena through close examination of the data. Open coding is conducted by separating data into discrete parts, and closely examining and comparing these discrete parts for similarities and differences, in addition to asking questions in regards to the phenomena as reflected in the data (Corbin and Strauss, 1990) and comparing them with the purpose of determining the underlying uniformity and its varying conditions (Glaser, 1978). Events, activities, objects and actions/interactions that are discovered to be conceptually similar in nature or connected in meaning are gathered under more abstract concepts that are referred to as "categories" (Strauss and Corbin, 1998).

Axial coding consists of the process of re-connecting the data that was separated by open coding, in new ways through establishing relationships between categories and their subcategories. This is enacted by fitting the pieces of the data 'puzzle' – which were fractured during open coding - together, considering that each piece (e.g., category, and subcategory) has its unique place in the overall explanatory scheme. When re-connecting the data or building the "puzzle", the first attempts of the analyst are often trial and error, however, as the analyst becomes more theoretically sensitive, making the fit between conceptual indicator and category becomes simpler. This process is referred to as "axial" since coding occurs around the axis of a category and links the categories at the level of properties and dimensions (Strauss and Corbin, 1998). Axial codes normally consist of categories that describe the open codes and during the process, coding and comparing the concepts to more incidents continue (Glaser, 1978). Comparison allows the researcher to identify the variations
Research process

in the patterns that are discovered in data; and the purpose of data coding at this level is to promote the data to higher levels of abstraction (Hutchinson, 1988).

**Selective coding** has the purpose of integrating and refining the categories into a theory that accounts for the phenomenon being investigated (Darke et al., 1998). In addition, it aims to validate the statements of relationships among concepts and complete any categories that require further refinement. In selective coding, data from many cases are reduced into concepts and sets of relational statements, that will be utilised to understand the phenomenon under investigation (Strauss and Corbin, 1998).

4.6.1.2 Memo writing

Memo writing, which is a device to represent the relationship between concepts, is a critical way of maintaining records of analysis. Martin and Turner (1986) and Strauss (1987) provide a comprehensive discussion of the processes involved in memo-writing. Memos are written throughout the research process, beginning with the first interview, and they provide a means of ensuring the research is grounded in data. Furthermore, memo writing enables the generation and development of explanations of the emerging concepts, in addition to discerning some of the inter-relationships that exist between them. The memo advises the meaning behind a code and acts as the pivotal step for dividing the categories into components and explaining the codes. According to Glaser (1978), theoretical memo writing is the fundamental phase in the theory generation process and is defined as "the theorising write-up of ideas about codes and their relationships as they strike the analyst while coding. A memo can be a sentence, a paragraph or a few pages. It exhausts the analyst momentary ideation based on data with perhaps a little conceptual elaboration". Memos do not only act as a data reporting tool and rather, combine various pieces of data into a recognisable cluster, which is frequently utilised to demonstrate that those data are instances of a general concept.

During analysis, memo writing is one of the most valuable and powerful sense-making tools that assists the analyst to transition from empirical data to conceptual level, by refining and expanding codes further and developing core categories and presenting their relationships, while building towards a more integrated understanding of events, processes, and interactions in the case. Memoing assists in establishing the core category around which the other categories integrate and hence provides the opportunity for the integration of the theory according to the emergent perspective of investigation and by defining its limits. As explained by Glaser (1978), the core category must be grounded in the data and "It must be central, i.e., related to as
many other categories and their properties as possible, and account for a large portion of the variation in a pattern of behaviour” (Glaser, 1978). Memos are a means of promptly capturing thoughts that occur during various phases (such as data collection, data reduction, data display, drawing conclusions and final reporting) by saturating dimensions of the core categories that have emerged through coding, and continuously generating open questions for further coding and data collection. At the end of the process, memos must be organised and integrated, which can simply be satisfied by grouping those memos that explain the same category, in order to clarify their dimensions and to distinguish them from other categories.

4.6.1.3 Theoretical sampling and comparison

According to Glaser and Strauss (1967), constant comparison and theoretical sampling are two analytical processes that contribute to raising categories to conceptual categories, and are both conducted through a process Glaser (1978) refers to as theoretical sampling and the selective sampling of the literature. Fundamentally, the researcher is required to challenge the conceptual categories with more data in order to define them vigilantly, explain their properties, explicate their causes, establish the conditions under which they operate, and determine their consequences.

The constant comparison is essential to the data analysis process in generating grounded theory. The aim of focused coding is to construct and elucidate a category by examining all the data it covers and variations from it. During the process of focused coding, the researcher applies a limited set of codes - that were established in the initial phase - to large amounts of data and subsequently assigns them to clusters or categories to which they have a clear fit. This process of comparison - where small parts of data are compared with other data and coded data is continuously confronted with new data for verification purposes - is named the constant comparative method by Glaser (1978). According to Huchinson (1988) "Comparative analysis forces the researcher to 'tease out' the emerging category by searching for its structure, temporality, cause, context, dimensions, consequences and its relationship to other categories". In addition, comparing the data categories and constructs that emerge between various groups of participants in the study, is appropriate and desirable because the purpose of the process of constant comparison is to generate a comprehensive and detailed theory. Fundamentally, it enables the researcher to transition quickly from describing the specifics of a case to thinking more abstractly about what commonalities the various cases share and how they vary (Lawrence, 2002).
Theoretical sampling commences during the data collection stage of the study and includes searching the transcripts for emerging categories that appear noteworthy and characterise the narrative. According to Glaser (1978) "Theoretical sampling is the process of data collection for generating theory whereby the analyst jointly collects, codes, and analyses data and decides what data to collect next and where to find them, in order to develop theory as it emerges. This process of data collection is controlled by the emerging theory, whether substantive or formal" (Glaser, 1978). The core purpose of theoretical sampling is to enable the researcher to determine the properties of the core variable under study by collecting new data to verify, complete and extend conceptual categories. Theoretical sampling and constant comparison are iterative processes that while being fluid and flexible, ensure that the analysis is planned and well-grounded in data and not disorganised and random, which can lead the analyst to unproductive paths and away from the focus of study (Lawrence, 2002).

Continuously comparing concepts according to their properties in order to uncover their similarities and differences, provides researchers with the opportunity to densify categories, differentiate between them, and to specify their range of variability. Once the analyst has obtained a few categories, the purpose of sampling is to develop, densify and saturate these categories. The sensitivity that a researcher has developed to the emerging concepts, is also related to this, where the more sensitive a researcher is to the theoretical relevance of certain concepts, the higher the likelihood of recognising the indicators of those concepts in the data. Sensitivity typically increases throughout the research project and assists the researcher in deciding what concepts to search for and where indicators of them can be discovered.

Continuous theoretical sampling can be utilised to increase the depth of focus and to ensure consistency, because constructs are derived from the data and therefore it is critical that data are collected in a systematic way for each category (Strauss and Corbin, 1990). According to Galser (1978) "theoretical sampling is used as a way of checking on the emerging conceptual framework rather than being used for the verification of preconceived hypotheses". Subsequent to the development of focused codes into categories, the researcher must combine them together in developing a grounded theory. The emerged grounded categories, which are derived from the data, are the fundamental building blocks for the theoretical understanding of the area under investigation. The conceptual framework that is developed from the conceptual categories is tested by gathering data, that provides support (or not) for the framework hypotheses and reveals the relationship between the categories, which form the basis for the subsequent emergent theory (Lawrence, 2002).
A critical challenge for theory building is to determine when sampling is able to be terminated. Ideally, sampling should impede when theoretical saturation is reached (Glaser, 1992; Eisenhardt, 1989). Theoretical saturation is the point at which incremental learning is minimal because no new information is arising since all new data result in information that have been encountered before (Glaser and Strauss, 1967).

In general, when building theory, the common rule is to collect data until each category is saturated (Glaser, 1978; Glaser and Strauss, 1967) which means until (a) no new or relevant data appear to emerge regarding a category, (b) the category is well developed in terms of its properties and dimensions demonstrating variations and (c) the relationships among categories are well-established and validated. The theory will not be evenly developed and it will lack density and precision, unless a researcher collects data until all categories are saturated (Lawrence, 2002).

**Conclusion**

In this chapter major areas of study have been covered, the philosophical and practical aspects of methodology selection when embarking upon empirical research. The interpretive approach and the research methods chosen for this study were described. The chapter has presented the research design and the data sources that showed how they contributed to the achievement of the research objectives. The limitation of research methods chosen was discussed. It has also presented and discussed grounded theory as a practical tool for analysing qualitative data and the reasons for using grounded theory to analyse the case study data. It concluded with a description of the procedures involved in doing data analysis in grounded theory. This chapter has established the general rationale for the study. The next chapter presents the analysis and findings of the survey questionnaire phase of the research.
5 Field studies: case study analysis

Introduction

This chapter describes the case study phase of this research, which has the purpose of obtaining in-depth and holistic qualitative information on how and under what conditions data analytics can result in innovation in enterprises. Consequently, the case study focused on deep understanding of the factors that influence the success of creating innovative products and services through the use of data analytics in enterprises. The data in this chapter is collected from 12 separate field studies that were conducted within the general practices of interpretive case study (Zuboff, 1988; Orlikowski, 1991; Walsham, 1993). The case study included extensive interviewing of main participants (e.g. innovation managers or data analytics centre managers), in addition to the use of documentary evidence such as company reports. In Chapter 4, I discussed how the case study was designed and I provided details of the research methods. Table 5.1. illustrates a number of key details of the organisations that participated in the case study.

5.1 Case study analysis

I studied and subsequently analysed 12 of the world’s largest and most innovative enterprises and in this section I discuss the analysis of the case studies conducted in this research. This discussion is based on research conducted by Orlikowski (1993), Glaser and Strauss (1967), Eisenhardt (1989), Miles and Huberman (1984; 1994) and Strauss and Corbin (1990; 1998).

The data analysis process included identifying patterns in the case study data that comprised of issues, innovation activities or opinions, which were raised frequently across interviews. To detect similarities and compare differences, the data were analysed within each case and across cases, i.e. the initial concepts that emerged in the context of one case were subsequently compared, elaborated, and tested with other cases.
Table 5.1: List of organisations that participated in our case studies

<table>
<thead>
<tr>
<th>Organisation</th>
<th>Size (employees)</th>
<th>Market cap</th>
<th>Established</th>
<th>Interview with:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook</td>
<td>17,048</td>
<td>$404.92 B</td>
<td>2004</td>
<td>Engineering Director</td>
</tr>
<tr>
<td>Air BnB</td>
<td>2000</td>
<td>N/A (revenue 2.8B)</td>
<td>2008</td>
<td>General Manager for design</td>
</tr>
<tr>
<td>Atlassian</td>
<td>1760</td>
<td>$6.68 B</td>
<td>2002</td>
<td>Product manager</td>
</tr>
<tr>
<td>Apple</td>
<td>98,000</td>
<td>$735.25 B</td>
<td>1976</td>
<td>Engineering program manager-technology innovation.</td>
</tr>
<tr>
<td>Google</td>
<td>47,756</td>
<td>$596.90 B</td>
<td>1998</td>
<td>Manager of marketing innovation.</td>
</tr>
<tr>
<td>Xero</td>
<td>1,450</td>
<td>N/A (revenue $302m)</td>
<td>2006</td>
<td>Developer relations manager and innovation manager.</td>
</tr>
<tr>
<td>Twitter</td>
<td>3,900</td>
<td>$10.96 B</td>
<td>2006</td>
<td>Engineering manager - Engineering team leading core parts of Twitter.</td>
</tr>
<tr>
<td>Wells Fargo</td>
<td>268,800</td>
<td>$296.63 B</td>
<td>1852</td>
<td>Marketing and strategy manager.</td>
</tr>
<tr>
<td>Visa</td>
<td>11,300</td>
<td>$209.70 B</td>
<td>1958</td>
<td>Marketing analytics manager</td>
</tr>
<tr>
<td>Netflix</td>
<td>3,500</td>
<td>$62.09 B</td>
<td>1997</td>
<td>Product innovation manager.</td>
</tr>
<tr>
<td>Tesla Motors</td>
<td>30,000</td>
<td>$43.32 B</td>
<td>2003</td>
<td>Data analytics manager.</td>
</tr>
<tr>
<td>Walmart</td>
<td>2,200,000</td>
<td>$231 B</td>
<td>1962</td>
<td>Walmart Labs Cognitive data scientist</td>
</tr>
</tbody>
</table>

Following the descriptions set out by Glaser and Strauss (1967) and Eisenhardt (1989) on how to generate grounded theory, for the first case the iterative method of data collection, coding, and analysis was more open-ended and generative, with a focus on the development of concepts, properties, and relations. All the detailed records, documents and data generated through interviews were examined and coded by focusing on the factors that influence the success of data analytics projects resulting in innovative products and services that create significant business or customer value.

The data generated by each case was examined and categorised using the open coding technique (Strauss and Corbin, 1990; 1998), which relies on concepts that were
proposed by the data rather than imposed from outside. Once all the data were inspected, axial coding (Strauss and Corbin, 1990) was conducted where the concepts were organised by recurring themes that became the main candidates for a set of constant and common categories, which linked several associated concepts. Axial coding relies on a synthetic method of formulating connections between subcategories to construct a more comprehensive scheme.

With the aid of this proposed scheme, the case data were re-examined and re-coded in order to determine collections of categories and concepts that covered as much of the data as possible. This iterative examination generated a collection of broad categories and associated concepts that explained the significant conditions, events and experiences associated with the creation of innovative products and services using data analytics in the case of the first organisation. These initial concepts, guided the analysis of the remaining case studies, by permitting the process of data collection, coding, and analysis to be more targeted. Using the constant comparative analysis method (Glaser and Strauss, 1967), the first organisation’s experiences were systematically compared and contrasted with the case of the second organisation and therefore data from the second organisation was first sorted into the initial concepts generated by the first organisation’s data. It was evident however, that the initial concepts that were generated by the first organisation’s case, did not accommodate several of the findings that had emerged from the second organisation’s case. The attempt to accommodate for the second organisation’s experiences, led to several critical explanations and clarifications in the emerging theoretical framework and forced a reconsideration of a few of the first organisation’s experiences. This process was repeated for all remaining case data from all the studied organisations. This analysis also utilised Miles and Huberman’s (1984; 1994) method for across-site pattern comparison and clustering that involves matrix displays to compare main events, triggers, and outcomes (Lawrence, 2002).

Redefining the initial concepts to incorporate the experiences and considerations arising from the second organisation’s case, necessitated that I return to the first organisation’s case data, and re-sort and re-analyse them to include the more comprehensive concepts and the more complex relations that now constituted the framework. This ability to incorporate distinctive insights while conducting the study is one of the advantages of the grounded theory methodology and is an example of what Eisenhardt (1988) labels "controlled opportunism," where "researchers take advantage of the uniqueness of a specific case and the emergence of new themes to improve resultant theory" (Eisenhardt, 1988).
Once sufficient number of categories and associated concepts had been defined to explain what had been observed at all the organisation cases and no additional data was discovered to develop or add to the set of concepts and categories, i.e. theoretical saturation was achieved, the iteration between data and concepts was terminated. The resultant framework is empirically valid because it is able to not only account for the unique data of each organisation’s case but also generalise patterns across all the organisations’ cases (Eisenhardt, 1989). The core categories and subcategories that emerged from the analysis are displayed in Table 5.2 below.

It is important to be vigilant about the validity of the interpretations made (Miles and Huberman, 1984; Yin, 1989) and therefore I examined the emerging concepts across participants and multiple methods to check their representativeness. Triangulation across data collection methods (interviews, documentation review and literature review) further assisted in strengthening the emerging concepts. In addition, the constant comparative technique, also compels the confrontation of emerging explanations with possible alternative ones by continuously searching for negative or non-conforming evidence. Ultimately, analysing the interview transcripts from each organisation’s case study, generated categories that provided the basis for constructing a qualitatively rich narrative description of the factors and conditions that enable successful innovation in enterprises through data analytics. This was subsequently presented to the interviewees, who commented, made corrections, and elaborated on drafts of the findings and conceptual model.

5.1.1 Propositions that emerged from the data analysis

This study proposes several research propositions that are derived from the analysis of the case studies and reflect the complex and dynamic nature of cultivating and enabling innovation through analytics in large organisations. The general process for obtaining these propositions was as follows: The initial propositions from the first case were systematically compared with evidence from the second case, to determine whether there is evidence to support, amend or invalidate them and therefore the initial propositions became a vehicle for generalising to the remaining cases. The process was repeated to refine theoretical propositions by systematically comparing them with evidence from the other cases to assess how well or poorly they fitted with data.

The fundamental idea is that I constantly compared theory and data - where data that supported the emergent theory enhanced confidence in its validity, while data that did not support the theory, frequently resulted in refining and extending the theoretical model - and therefore I iterated towards a theory that fits the data closely.
Field studies: case study analysis

<table>
<thead>
<tr>
<th>Core Categories</th>
<th>Subcategories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance of the Problem</td>
<td>Understand the business challenges and problems.</td>
</tr>
<tr>
<td></td>
<td>Understand customers and focus on them.</td>
</tr>
<tr>
<td></td>
<td>Focus on ideas that result in 10x better.</td>
</tr>
<tr>
<td>People aspect</td>
<td>Work in teams.</td>
</tr>
<tr>
<td></td>
<td>Quality of people.</td>
</tr>
<tr>
<td></td>
<td>Multi-disciplinary teams.</td>
</tr>
<tr>
<td></td>
<td>Collaboration between teams.</td>
</tr>
<tr>
<td>Internal Organisational Culture</td>
<td>Collaboration and creativity</td>
</tr>
<tr>
<td></td>
<td>Innovation culture</td>
</tr>
<tr>
<td></td>
<td>Failure Accepting Culture</td>
</tr>
<tr>
<td></td>
<td>Existence of an Innovation Framework</td>
</tr>
<tr>
<td>Experimentation Factors</td>
<td>Prototyping</td>
</tr>
<tr>
<td></td>
<td>Iteration</td>
</tr>
<tr>
<td></td>
<td>Fast-paced execution</td>
</tr>
<tr>
<td>Managerial support</td>
<td>Manager’s as team members</td>
</tr>
<tr>
<td></td>
<td>Adequate allocation of resources</td>
</tr>
<tr>
<td></td>
<td>Managers understanding the limits and power of data.</td>
</tr>
<tr>
<td>Variety of Data</td>
<td>Multiple data sources.</td>
</tr>
<tr>
<td>Organisational Structure</td>
<td>Physical environment.</td>
</tr>
<tr>
<td></td>
<td>Separate small units.</td>
</tr>
</tbody>
</table>

Table 5.2: Core categories and subcategories that emerged from the data analysis.

(Eisenhardt, 1989). Generally, the propositions established in this case study, defined the preliminary categories that presented the factors that influence the success of data analytics efforts in resulting in innovative products and services. These proposed propositions are provided in table 5.3 below.
### Field studies: case study analysis

<table>
<thead>
<tr>
<th>Core Categories</th>
<th>Subcategories</th>
<th>Propositions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance of the Problem</td>
<td>Understand the business challenges and problems.</td>
<td>Directing the efforts of a big data project towards an actual business problem and challenge will influence the success of the analytics efforts creating value for the business.</td>
</tr>
<tr>
<td></td>
<td>Understand customers and focus on them.</td>
<td>Focusing on customers’ needs and pain points increases the probability of success of creating significant customer and business value.</td>
</tr>
<tr>
<td></td>
<td>Focus on ideas that result in 10x better.</td>
<td>Aiming for an order of magnitude improvement will force a total re-imagination of products and services that increase the probability of innovating.</td>
</tr>
<tr>
<td>People aspect</td>
<td>Work in teams.</td>
<td>Groups often make better decisions than individuals and therefore streamlined team work is an important factor for the success of innovation projects.</td>
</tr>
<tr>
<td></td>
<td>Quality of people.</td>
<td>People in teams influence the success of data analytics projects resulting in innovation.</td>
</tr>
<tr>
<td></td>
<td>Multi-disciplinary teams.</td>
<td>Having multi-disciplinary teams of curious, thoughtful, and motivated learners with a broad range of interests influences the success of data analytics projects resulting in innovative products and services.</td>
</tr>
<tr>
<td></td>
<td>Collaboration between teams.</td>
<td>Small teams are useful but they need to collaborate with each other to increase the probability of success.</td>
</tr>
<tr>
<td>Internal Organisational Culture</td>
<td>Collaboration and creativity</td>
<td>A culture that encourages collaboration and creativity influences the probability of success of data analytics projects.</td>
</tr>
<tr>
<td></td>
<td>Innovation culture</td>
<td>An innovation culture that is distributed across the whole organisation is necessary.</td>
</tr>
<tr>
<td></td>
<td>Failure Accepting Culture</td>
<td>A culture that accepts failure as an inevitable part of success increases the innovation ability.</td>
</tr>
<tr>
<td></td>
<td>Existence of an Innovation Framework</td>
<td>Establishing a framework improves the use of time and resources, increasing the innovation ability in a systematic way.</td>
</tr>
</tbody>
</table>
Field studies: case study analysis

<table>
<thead>
<tr>
<th>Experimentation Factors</th>
<th>Prototyping</th>
<th>Building and prototyping is an influencing factor of the innovation ability by providing fast and effective feedback.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Iteration</td>
<td>Idea has value when its executed. Executing fast, learning and then iterating is an influencing factor in ultimately succeeding in innovation.</td>
</tr>
<tr>
<td></td>
<td>Fast-paced execution</td>
<td>Moving fast, implementing, learning, iterating, and if necessary killing a project fast are important factors that influence success of achieving innovation.</td>
</tr>
<tr>
<td>Managerial support</td>
<td>Manager’s as team members</td>
<td>Changing the role of managers from judges to active participants in teams influences the probability of success of achieving innovation.</td>
</tr>
<tr>
<td></td>
<td>Adequate allocation of resources</td>
<td>Organisational resources influences the rate of importance given to innovative thinking, while the lack of it inhibits dedication and ultimate success.</td>
</tr>
<tr>
<td>Variety of Data</td>
<td>Multiple data sources.</td>
<td>Power of big data can mainly be realised if multiple sources of data including external data are used.</td>
</tr>
<tr>
<td>Organisational Structure</td>
<td>Physical environment.</td>
<td>The physical environment influences the creativity of team members and how they think, consequently affecting the success of data analytics projects resulting in innovative products and services.</td>
</tr>
<tr>
<td></td>
<td>Separate small units.</td>
<td>Small teams that are devoted to finding and implementing new innovative ideas, can influence the success of innovation efforts.</td>
</tr>
</tbody>
</table>

Table 5.3: Propositions that emerged from data analysis

5.2 Results of case study analysis

This section presents the results of the organisations’ experiences using data analytics to create innovative products and services and cultivate innovation in general. The grounded theory technique, discussed in detail in Chapter 4 (Glaser and Strauss, 1967; Strauss and Corbin, 1998), was used to analyse the data derived from the case studies. The case study results were utilised to develop a theoretical model for conceptualising the organisational factors that are required for the success of data analytics driven innovation. In the following sections, the results of the multiple case studies will be discussed in terms of the categories that emerged from the
grounded theory analysis process, in order to recognise and provide substance for each category.

Table 5.3 depicts the success factors that emerged from the innovative organisations’ experiences with the use of data analytics for innovation and it introduces the categories and concepts that emerged as significant from the data. No claim is made that the factors and categories presented here are exhaustive and indeed further organisational studies of the use of data analytics for innovation, should enhance or modify the ideas presented here. (Orlikowski, 1993). The categories constituting the factors that influence or inhibit the ability of an organisations to achieve innovation through analytics are discussed below.

5.2.1 Importance of the problem

The case results indicate that across all enterprises studied, the most critical factor that influences the success of analytics projects in developing innovation, is focusing on solving a genuine problem that has the potential to offer significant value to the customers and/or business.

Visa Reports: “The challenge is that big data analytics must be driven by the right use case to permit businesses to scale analytics to make money or reduce costs.”

Amazon Reports: “if acting in isolation, big data and data scientists do not possess a magic formula that is capable of radically transforming businesses or changing the world. The act of solving problems is undoubtedly different from accumulating a tremendously large data set or allowing several data geeks to have free reign on data. . . . In our experience, what makes data valuable is problem solving, and solving problems with data requires a diversity of thought and a methodology that balances number crunching with thoughtful design to solve targeted and specific problems.”

The case studies provide support for the importance of taking a focused approach in data analytics projects and aiming to solve a specific problem in these projects. This is consistent with prior research that was discussed in section 2.3 that indicated the most effective big data and analytics strategies identify business requirements first and then leverage the existing infrastructure, data sources and analytics to support the business opportunity (IBM, 2013). Existing literature proposes that the power of data does not eliminate the need for vision and human insights; and being successful requires thinking creatively and proposing truly novel offerings (McAfee and Brynjolfsson, 2012).
Furthermore, Stubbs (2014) argues that organisations that succeed in obtaining tangible value from data analytics have discovered that analytics is concerned with value and outcomes and therefore these firms bring together the business, IT, and the analysts into a coordinated team focused on value creation. Rather than focusing on insights, such organisations attempt to directly link their analysis to measurable results. Although these organisations are still concerned with the sophistication of analysis and the efficiency of algorithms, their primary goal is to drive significant outcomes and it is understood that insights without action are valueless.

There are three major ways in which the innovative organisations in our study increase the probability of success of discovering or devising a problem or opportunity that will add significant business or customer value. These approaches are understanding the business challenges and problems, understanding the customer and focusing on them, and striving for solutions that result in 10x better offerings.

5.2.1.1 Understand the business challenges and problems

The case results demonstrate that understanding the pain points of the business and various business units, is a significant factor contributing to the increased probability of identifying and devising a problem or opportunity that can result in significant value.

Facebook reports: “For us the most significant issue is to discover what the deepest problem the business has is and then proceeding to using whatever technology is required to solve it. It’s not the other way around. We find the problem, we solve it, and then repeat.”

Google reports: “you can’t create something innovative using analytics, unless you get the foundation right and really understand the user, the insights, and the strategic issue that you have interrogated”

Amazon reports: “Considerate design to solve targeted problems is essential to the success of analytics. But I don’t think that means to have a significant number of PhDs with deep knowledge on every topic you can think of… what is important is to find what the business problem is and what is keeping them up at night; and it’s not easy to find that, you have to be creative about finding it.”

Similar results have been reported in literature, where the importance of solving business challenges and problems with analytics has been emphasised. For example, in their study, Lavalle et al. (2011) reported that performing analytics without strategic business direction is not appropriate because the efforts are likely to stall,
Field studies: case study analysis

which not only will waste resources but also it risks creating widespread skepticism about the real value of analytics. According to Lavalle et al. (2011) analytics aligned to a significant organisational challenge however, makes it simpler to overcome a wide range of obstacles that are significantly important to the success of analytics projects and are required to be addressed in a timely manner, such as executive sponsorship of analytics projects, data quality and access, governance, skills and culture. When overtaken by the momentum of a single significant idea and potentially game-changing insight, obstacles like these have a greater probability of getting swept into the wake of change rather than drowning the effort.

Visa further emphasises the importance of the problem and reports: “It is extremely complex and resource-intensive to build an analytics system, particularly one that involves Big Data, and therefore many organizations hope the software they deploy will be a magic bullet that instantly does it all for them. This is obviously not correct. While software does assist and sometimes in a significant way, Big Data analytics is only as good as the problem that is trying to be solved. Data being analyzed and the analytical skills of those using the tools are also critical but the most important factor is what you are actually trying to solve.”

Airbnb similarly believes that performing effective data analytics is not solely concerned with the technology aspects and rather understanding the business domain plays a critical role: “A consequence of the misconception that technology can be a solution to everything, is the belief that all an organisation requires is analysts that can implement the analytics software. If you are to generate business value, an effective analytics program whether with big data or small data, has to incorporate extensive business and industry knowledge into both the system design stage and ongoing operations.”

Likewise, Walls Fargo believes that analytics is required to be driven by business needs, and reports: “Our Data Science team frequently begins Big Data-driven analytical activities in any division or part of the business where there is an indication that it can offer value. Identifying where analytics resources can most effectively be used is challenging, and for us it involves lining up business use cases with technological capabilities.”

Similarly Google reports: “Since the establishment of our Data Innovation program, we have executed hundreds of proof-of-concepts and use cases that were all validated against satisfying certain business requirements. Our focus is on obtaining actionable results that are connected to very specific metric-based outcomes. Therefore, every time a potential use for analytics is identified, we assess it in regards to its
expected benefits and opportunity cost. We take many different factors into account, which is why not all of these opportunities will get selected, and even then sometimes half way through the project we may think that there are better ways to achieve this and moving to Big Data just because it is Big Data may not be suitable.”

5.2.1.2 Understand customers and focus on them

The case results reveal that understanding customer pain points, needs and wishes significantly increase the probability of creating innovative products and services using data analytics that generate significant business and customer value. The enterprises in our study believed that understanding both mainstream and emergent customers is a critical factor in the success of data analytics projects resulting in significant innovation. This is in line with innovation research discussed in Chapter 3.4.4 where it was argued that insights about user needs (especially emerging customer segments) are a superior starting point to technological development or visions within the firm (Veryzer and Borja deMozota, 2005; Jansen and Dankbaar, 2008; Hamel and Prahalad, 1994; Verganti, 2008; Christensen 1997; Govindarajan et al., 2011).

Wells Fargo reports: “To be successful in analytics, it’s not about disruption or being “cool”, it’s about listening to the customer in terms of what they require us to provide for them and what their deepest needs are.”

Similarly Google reports: “From a strategic point of view, I believe it is crucial to understand the users in particular, and spend a significant amount of time getting insights by for example going deep into the quantitative and qualitative user data, and attempting to “champion” them. As a strategic starting point, this is where we focus on aggressively, and try to get to the heart of why we’re doing analytics because unless we get this foundation right and really understand the user, the insights, and the strategic issue, it is difficult to create something innovative.”

Further to these insights, Wells Fargo has been focusing on customer centric innovation for nearly two decades. In our interview Wells Fargo reports: “many years ago we realised that there is a significant opportunity to use the internet to create new value that didn’t exist in the paper-based world. We understood that we were able to build a banking business around customers, not products; which meant clients could spend less time dealing with banking and more time running their businesses. Consequently, we immersed ourselves in the internet culture and ultimately used ethnography techniques, which at that time were being used frequently in consumer
products companies like Procter & Gamble, but were not used in the banking industry at all.

Ethnography techniques were significantly suitable for us as it enabled us to get out and observe customers as they navigated the typically confusing banking routines, instead of having executives brainstorm in the corporate offices. This way by uncovering problems and experienced frustrations first-hand, we were able to devise solutions.

Today, we go into customers’ offices 30 times a year and watch them bank, in addition to having customer councils to advise where the pain points are and how service can be improved and therefore at Wells merely devise financial solutions and instead we operate in product design.

Twice a year, our team uses the ethnography studies, customer councils and insights that are collected from internal metrics to identify ten priorities that will then provide us with directions to develop, test and deploy new services every 90 days. This has been done for more than fourteen years.”

Amazon similarly reports on the importance of understanding customer needs in their analytics endeavours: “What’s critical for us to consider is that at the end of every technology that we develop, there is a customer. Therefore, it is crucial for us to listen to our customers, pilot, test, and ask them, not in terms of what we want to provide to them, but so we can be responsive to their needs.”

These results are consistent with innovation theory literature that places a significant emphasis on the importance of customers discussed in section 3.4.4. Dennings (2014) argued that the most effective way for large organisations to cope with big bang disruption is understanding, anticipating, and meeting the needs, wants and whims of customers that are well informed, empowered and interactive. Dennings (2014) added that corporations must learn to delight customers through continuous, disciplined, and transformational innovation. On the other hand Govindarajan et al. (2011) similarly to Christensen (1997) believed that emerging customers instead of mainstream customers should be focused on and argued that an orientation to mainstream customers was significantly and negatively related with radical innovation and that an orientation to small, emerging customer segments was positively related to radical innovation success.

In this context Amazon reports: “At Amazon we have a common line of thinking: “start with the customer and work backwards. This is effective because when you work backwards, you begin with the customer and their needs and problems, which
is in contrast with what some large companies do, which is devising ideas, building products, and then testing to see if customers like it. Bezos our CEO in fact famously said: “if we can arrange things in such a way that our interests are aligned with our customers, then in the long term that will work out really well for customers and it will work out really well for Amazon”.

5.2.1.3 Focus on ideas that result in 10x better

If the aim of an organisation is to achieve truly radical and discontinuous innovations that are possible using data, small improvements are not what they should be striving for as they result only in incremental innovations. The case results show that attempting to improve an existing product or service by at least an order of magnitude, necessitates the total re-imagination of how a problem is approached and thus increases the probability of creating radical innovations.

Google reports: “The “10x thinking” is at the heart of how we innovate at Google and that extends to innovation using analytics. The idea behind this is simply that we believe true innovation occurs when you attempt to improve something by 10 times rather than by 10%....This is especially the guiding inspiration for engineers at Google[x]—(the division of Google that focuses on producing major technological advances, from glucose-monitoring contact lenses, balloons that deliver Internet access to remote areas of the world, to self-driving cars). . . . For a 10% goal we must rethink an idea entirely and go beyond existing models, which forces us to entirely re-imagine how we approach the idea.”

Amazon reports: “No one is going to be willing to change the status-quo for anything less than an order of magnitude. Anything less, you are wasting your time.”

5.2.2 People aspect

Unsurprisingly, the case results indicate that one of the crucial factors that influence the success of data driven innovations in organisations, is the people component. Three categories that emerged from the analysis include quality of people, team work, and collaboration between teams.

5.2.2.1 Quality of people

All enterprises in our case studies indicated that without the appropriate people with adequate skill sets and behaviours, an innovation project has a low probability of success. The emotional, social and intellectual ‘assets’ of people, that include mo-
tivation, style, knowledge, skills and experience, all play crucial roles in developing innovation through data analytics projects.

Google reports: “We look for people who are great at a broad range of attributes, love enormous challenges and welcome change. The most vital principal we have at Google is the people-centric approach. In our analytics team, we look for a wide range of people such as entrepreneurs, olympic winners, extremely intelligent and dedicated people. We design the company and the leadership, after those kinds of talents to ensure that we can attract them and bring them in.”

Twitter reported: “Best thing we have ever done to cultivate innovation is hiring the best people. We treat our employees as our second customers.”

The case results are also compatible with the findings of prior research discussed in section 3.4.1, that indicate the importance of employees in teams for successful innovations. The criticality of the team members and its leader has also been shown for the success of the front end of innovation in the literature. Two meta-analyses of new product teams, by Hulsheger et al. (2009) and by Sivasubramaniam et al. (2012), revealed that the effectiveness of the team leader, the team’s overall cognitive ability and experience, and how well the team worked together, had a significant relationship to the innovation constructs of the studies surveyed in the analyses. Similarly, in their study Koen et al. (2014) presented the significance of teams and collaboration for the success of front end of innovation and provided a summary of their necessary attributes for success in three constructs:

- **Effective Teams**: Team members are passionately committed to their front-end projects and spend time and effort on them beyond that required by their job.

- **Team Leadership**: Team leaders have established credibility and recognized leadership experience.

- **Communities of Practice (CoPs)**: The company supports CoPs, provides them with a budget, and has a coordinator who dedicates at least 25 percent of his or her time to the community.

Amazon also stresses the importance of people for enabling innovation and reports: “At our team we think innovation is a point of view and therefore it is necessary to actually select people that are a part of the company who want to innovate and explore. Being a part of a pioneering and an exploring team is not for everyone and therefore we need to attract like-minded people who desire to invent. In our
Field studies: case study analysis

analytics team we look for people who wake up in the morning and think about how we can work backwards from customers and build a great service or a great product. That is a critical element for enabling invention and innovation. . . . . . . if your team has the right kind of people that like to invent and are pro-change, then everything they see as they move about the world they think about how it could be improved.”

In addition, Visa believes in the importance of analytical skills in the development of innovation and reports: “many enterprises underestimate the extent of the analytical skills that are required. If Big Data analysis was only about constructing reports and dashboards, then their existing BI expertise would have been adequate but Big Data analytics involves more advanced processes (such as data mining and predictive analytics). Therefore, analytics professionals with statistical, actuarial, and other sophisticated skills, are required.”

5.2.2.2 Team work

Working in teams to enable innovation through data analytics, is another important factor that was revealed in our case studies. Interdisciplinary teams are associated with faster development rates (Stalk and Hout, 1990; Eisenhardt and Tabrizi, 1995) and enhanced performance (Brown and Eisenhardt, 1995). The use of these types of teams has also been connected to enhancements to organisational efficiency and quality (Applebaum and Batt, 1994), and financial performance (Macy and Izumi, 1993). In addition, working in teams enable the flexibility required to effectively and efficiently respond to the continuously changing demands of the organisation’s environment (Zaccaro et al., 2001) and facilitate the development of the various human dimensions involved during the strategy-related activities that include cognitive, emotional and social (Roos and Victor, 1999; Kerr et al., 2012).

Google reports: “There is a significant amount of evidence that indicates that groups make better decisions than individuals, especially when the groups are selected to be among the most intelligent and most interesting people. For analytics we certainly try to improve the probability of innovation by having a diverse group of people work together, and let the great minds work collectively to create innovative solutions for hard problems.”

Similarly, Visa reports: “To achieve innovation through analytics, it is critical to assemble an interdisciplinary project team to work collaboratively and then keep analysing how the group interacts in order to select a makeup and methods that best promote individual thinking rather than groupthink. It is important to have a team that fosters multiple perspectives, quick production and fluid communication
because such teams can take advantage of diversity, which is important to the innovation process whether it's through analytics or not. Also, such groups are more likely to thrive in organizations that nurture “cultures of innovation” and therefore you are required to reward risk taking, encourage people to mix, support play and new ideas, don’t demonize failure, and don’t overemphasize regulations or efficiency.

5.2.2.3 Collaboration between teams

Another innovation success factor derived from our case studies was the importance of collaboration between different groups within the organisation. While having small teams that are dedicated to innovation is important, having collaboration between these teams is also critical for cultivating innovation through data analytics.

Atlassian reports: “It’s a vital part of [maintaining a] competitive barrier to have resilient, scalable, differentiated, and global products, which means they can no longer be built by two people. Therefore, for our analytics teams, while we do recognise that innovation is driven from small teams and we have such small teams, we also encourage them to talk to each other and collaborate.”

Google reports: “One of the things that we’ve vigorously attempted to avoid at Google is the sort of divisional structure and the business unit structure that prevents collaboration across units. While this is a challenging task and I understand why there is a desire to construct business units and have their separate presidents, by doing that you to a large portion eliminate the informal ties that, in an open culture, drive so much collaboration. If you build a strong shared value culture, then people in the company have a deep understanding of the values of the firm, and therefore they should be able to self-organize to work on the problems that are most interesting to them.”

5.2.3 Internal organisational culture

The case evidence reveal that the internal organisational culture affects the innovativeness and success of the analytics endeavours significantly. In fact, the enterprises in our study emphasised that introducing a change in cultural philosophy is one of the most important chores associated with leveraging analytics. Stubbs (2014) argues that deriving business and customer value from data analytics is impossible unless the organisation has an adequate innovation supporting culture in place. Culture is crucial because value can only be obtained from analytics if people are able to work together in harmony and agree on what is to be achieved and what the goals are. Furthermore, it is necessary to construct a new culture of technology
professionals who also have significant business skills. Developing such culture is dependent on many factors, such as ensuring that the teams are educated about the business culture, and also emphasizing measurements and results. According to Ohlhorst (2012), initiating this at the top and with leaders, proves to be one of the most effective means of transforming an IT-centered culture into an internal business culture that thrives on advanced data analytics technology and fact-based decision making.

The categories that emerged from the case analysis include innovation culture, failure accepting culture and innovation frameworks.

5.2.3.1 Innovation culture

In our case studies, having a strong innovation culture that cultivates innovation, was exposed as a significant factor influencing the success of data analytics projects in creating significant business and customer value.

Atlassian reports: “Innovation culture means everything in terms of your analytics success because products come and go but it is our culture that stays. We are not scared to disrupt ourselves and therefore we are dynamic in a sense that we are actively seeking what is occurring in terms of trends, and then bring those insights in, and reallocate their recourses. But what allows us to take advantage of all this is our culture, which stays constant.”

Apple reports: “Our culture is distributed across the entire organisation and all the people that work at Apple. Our innovation culture is critical for our success”.

Facebook reports: “Our innovation culture is extremely important to all aspects of what we do at Facebook including analytics. We are an incumbent rather than an upstart, which means that if it we are not careful, we could ultimately be in the same position as numerous other tech companies that began with a powerful culture but lost it along the way and became just a bureaucratic and large company that is not able to move fast . . . . Maintaining our culture doesn’t get simpler as we grow, but we don’t believe that we should keep the culture exactly the same. We have certain values but how we do things to achieve them changes over time. Initially in the analytics teams just like the entire company we were following: ‘We can tolerate mistakes. ‘Move fast and break things’. But as we scaled that was not feasible anymore because we were making so many mistakes that we were being slowed down by them so we changed our strategy to ‘Move fast with stable infrastructure.’”
This result is consistent with prior research that was reviewed in section 3.4.2.4. Many scholars believe that a well-established culture and process of innovation inside the firm is a fundamental factor influencing the rate of creation and commercialisation of innovation outcomes (Myers and Marquis, 1969; Xu et al., 2007). Organisational culture and integral elements of culture, such as entrepreneurship, risk-taking, flexibility and creativity, are particularly important and must be preserved and valued in order to develop innovation in general and disruptive innovation in particular (Govindarajan and Kopalle, 2006; Murase, 2003). Moreover, the internal culture within the firm plays a critical role in inspiring innovation and providing individuals with adequate freedom to make mistakes, which assists in creating more opportunities for serendipity and valuable learning (Peebles, 2003). According to Teece (2010), organisational culture is not “only important for creating value (for the firm in the market) but of capturing it as well”.

Drucker (1993) proposed several critical factors to create entrepreneurial and innovation culture inside the organisation and overcome the resistance to an innovative environment. Table 5.4 outlines these factors:

| Policy | • Innovation preserves and perpetuates organisation.  
|        | • Innovation needs and its timeframe.  
|        | • Innovation plan with specific objectives and:  
|        |   • Systematic policy of abandoning obsolete things  
|        |   • Free people to innovate and seek new things  
|        |   • Allocate financial resources  
|        |   • Requirements, areas, and timeframe.  
| Managerial Practice | • Focus managerial vision on opportunity (report problem vs. opportunity).  
|                    | • Generate entrepreneurial spirit through entire management group.  
|                    | • Top management meet with junior personnel.  
| Innovation Performance Measurement | • Feedback from results to expectation in innovative project.  
|                                  | • Systematic review and valuation (objectives vs. performance).  
| Framework | • Includes structure, staffing, compensation, incentives, and rewards.  
|          | • People to be entrepreneurial and innovative rewarded not penalised.  
|          | • Separate new unit (innovative project) from old unit.  
|          | • Assign a special manager for new unit.  
|          | • Separate and apply different measurement for return-on-investment analysis.  
|          | • Accountability.  
| Don’ts | • Mix managerial units and entrepreneurial units.  
|       | • Diversify innovation, focus on similar business field.  
|       | • Acquire small entrepreneurial venture.  

Table 5.4: Entrepreneurial and innovation culture factors (Drucker, 1993)

In addition, Kenny and Reedy (2006) proposed four types of attitudes that are required to exist in an organisation’s innovative culture in order to succeed and
flourish, which include: risk-taking management, members’ participation, creativity stimulation, and sharing responsibility. An organisational culture that is based on innovation, must be able to deal with goal commitment, exemplary behaviour, team work approach, client orientation, and continuous improvement; and without a shared innovation-oriented culture it is difficult to compete (Deshpande et al., 1993).

5.2.3.2 Failure accepting culture

Having a culture that promotes experimentation and understands that failure is an inevitable part of ultimate success, is an extremely important factor for the success of data analytics projects in value creation.

Google reports: “Our culture at Google - which of course extends to the analytics teams- is failure accepting and reflects what Larry (CEO of Google) says: “It’s okay to fail if you fail fast and you learn something from it, and then you move on.” The failure accepting aspect is ingrained in our culture, and is at the heart of what we do when using analytics for innovation. We have had many failures and also if we recognise through a project that it might not be working, we can adapt or change or very quickly learn and move on. I think the risk is when you take too long to establish if something is not working/ not adding value and you don’t learn from it. That’s when it becomes a risk but you can mitigate that when you share the learnings and look at failure as a tool for learning that will ultimately lead to creating a tremendously valuable product and service.”

Facebook reports: “We cannot say that we have made it so let’s get comfortable and stop innovating. If we do that, in a year we will be gone and people will say “remember Facebook? I used to use that a lot.” Therefore we have to change constantly, and an inevitable part of seeking change and innovation is failure. The worst thing you can do for change is to create fear that if you try and subsequently fail, we will have a blame culture.”

The case results are supported by other findings in literature. The Probe and Learn Process by Lynn et al. (1996) suggests that permitting making mistakes during innovation activities is essential to learning and creating new knowledge. This way of thinking, initiates a change of the corporate environment and improves the innovation culture.

In addition, in their exploratory study, Brem and Voigt (2007) argue that innovation culture is a highly relevant aspect in view of personal motivation. An example response from one of their interview partners was: “if an idea gets through into a
successful innovation, no one will notice. But if it fails, then you will be blamed for that. So finally, you have no chance to win something”. Based on their study, Brem and Voigt (2007) concluded that a firm must motivate its employees and have a culture in place where failure is accepted as an inevitable part of innovation, otherwise no above-average results can be expected.

In our study, Google further emphasised the importance of a failure accepting culture and reports: “People remember your hits more than your misses that’s why Google is known for YouTube, not Google Video Player. In our analytics projects specially when focus is on innovation, we believe that It is okay to fail as long as we learn from our mistakes and correct them fast. Knowing that it’s okay to fail can encourage people to take risks and that’s essential for us because the tech industry is so dynamic that if we stop taking risks we will get left behind and lose our competitive edge.

For example two projects that we had: AdSense and Google Answers, were both radically new for the company but while AdSense is now a multi-billion-dollar business, Google Answers (which lets users post questions and pay an expert for the answer) was eliminated after four years. We could think of Google Answers as a failure but we learned immensely in that time, and we were able to apply the knowledge acquired to the development of future products. If we had a culture that would make of afraid of failure and wouldn’t encourage us to take risks, we never would have tried Google Answers or AdSense, and would have missed an opportunity with each one.”

Similarly, Amazon reports: “For being able to achieve innovation through analytics a fundamental success factor is to have a willingness to fail or sometimes be misunderstood for long periods of time. If something is done in a new way, people are initially going to misunderstand it relative to the traditional approach. If we were to try to not fail and always ensure that we will succeed, we would never be able to innovate using analytics and so ultimately we would not succeed.”

5.2.3.3 Innovation framework

The influence of having an organisation wide framework for innovation was also indicated in our case studies as an influencing factor for the success of data analytics projects. In order to drive repeatability and process efficiency, the organisations in our study have typically standardized their processes to varying degrees and have minimized transaction costs through effective use of workflows (Stubbs, 2014). It is believed that risks associated with creating innovative products and services using data analytics, can be reduced by following repeatable processes that help the
organization to change and innovate. Many of the organisations in our study are striving not only for outcomes, but also for repeatability and reuse as economies of scale and scope become real and provide cost advantages over competitors.

Airbnb reports: “It might be tempting to dive promptly into problem-solving using analytics without establishing an appropriate framework, but that is a mistake because it can lead to productivity losses and spending a significant amount of time and resources on products that should never had been designed in the first place.”

Wells Fargo similarly believes that it is critical to approach innovation in a systematic way, and reports: “The difficult challenge of using data analytics for most organisations is how to turn all that rhetoric into a revenue growing and value creating reality. And the aim is to not just make incremental changes to existing products or services, or pursue a one-off success, but to produce a continuous stream of breakthrough innovations that compound over time to shape a lasting competitive advantage. You are required to work systematically on the innovation management challenge and confront it in a broad-based and exceedingly systemic way.”

Likewise Google reports: “Our objective is to be a systematic innovator at scale, which means systemizing our approaches so we can enable groups to innovate. With our analytics projects we don’t necessarily know which one will succeed each month but what we know is that we can satisfy the portfolio theory i.e. we have an adequate number of innovation projects that while not all of them will succeed, there will be some of them that will generate innovation. And our systematic ways to innovate, help with ensuring that we always have a pipeline of potential new innovations.”

Visa reports: “For our analytics team, we are driven by the continuous desire and necessity to improve. And this desire for recurrent efficiency and efficacy gains is at the core of everything we do; we understand both the value of data and the need to act upon it and we have gained the experience to execute it. At Visa we no longer have the challenge of selling the value of analytics and instead we are concerned with making it pervasive and so defined processes are gradually becoming the norm and we also attempt to use analytics to support even micro-decisions.”

These findings are consistent with past research that was reviewed in section 3.4.3. O’connor and Ayers (2005) reported that since many organizations struggle with increasing their innovativeness, they often view innovation as synonymous with creativity and “coming up with great ideas” (O’Connor and Ayers, 2005). Hence, they turn to solutions such as new ideation methods, idea jams, and involving users in ideation or web-based crowd-sourcing. However, a single focus on idea and concept generation is disputed, and it is argued that more attention is required to be paid to
implementation and a holistic view of the innovation process. In fact, because the development of new products and services is an immensely complex procedure, in large organisations product and service innovations typically need to be systematically prepared, realized and implemented in order to be successful and that is done through an innovation framework (Stockmeyer, 2001), (Hauschildt et al., 2011) and Pleschak et al., (1996).

As discussed in Chapter 3.4, in situations where the market uncertainty is low and the technological uncertainty is high, or vice versa, the focus of an innovation framework should be on activities that exploit the existing knowledge and reduce the residual risk. In such cases, separating the ideation process to strictly sequential phases will not meet the requirements of reducing and minimising technical or market uncertainty. According to Verworn et al. (2007) in these situations, a learning-based strategy and an iterative procedure is required. Eisenhardt and Tabrizi (1995) believed, in uncertain conditions, an ‘experimental model’ can be a superior fit, allowing for improvisation and flexibility, and where learning is accelerated through iterations and testing, in combination with strong motivation and leadership.

Similarly, for radical innovation that is the most extreme case of innovation and seeks both new markets and technologies, all areas and functions must proceed gradually through extensive processes of learning and experience. For this purpose, the process must have the necessary openness to guarantee iterations and to make the integration of feedback possible at the right time (Verworn et al., 2007).

5.2.4 Experimentation factors

Taking an experimental approach to data analytics for innovation was another important factor stressed by the organisations in our study. This is consistent with the literature results discussed in section 3.4.3.2. It was revealed that for discontinuous innovations where uncertainty is high, an ‘experimental model’ for innovation and especially for the front end of innovation, can be an appropriate fit, allowing for improvisation and flexibility, and also for learning to be accelerated through iterations and testing.

Amazon reports: “If you are attempting to achieve innovation through data analytics on a regular basis, you must increase your rate of experimentation. You need to determine and discover how to organise the systems, people, assets, etc. to increase your rate of experimentation. Not all experimentation will result in value and therefore it is important to be consistent.”
Field studies: case study analysis

The categories that emerged from the analysis include prototyping, iterations, and face-paced executions.

5.2.4.1 Prototyping

The organisations in our study all emphasised the importance of prototyping as a means of identifying, testing and validating the assumptions of critical importance in data analytics endeavours as early as possible. Their use of prototyping is primarily about experimentation as an early mechanism to mitigate the risk of considerably wasting resources later by developing a ‘solution’ that is not desired or needed, which might lead to costly and time-consuming activities to solve the situation.

Facebook reports: “Our most important principle is to facilitate learning by making. We don’t have a review board that our analysts, designers and engineers present to using PowerPoint slides, instead we are without a doubt a build and prototype culture and we believe ideas presented on slides don’t speak on a personal level and just don’t stick. So prototyping and delivering a real tangible solution as early as possible is of utmost importance in our analytics projects.”

Prototypes can take various forms and are used to facilitate learning and obtain necessary feedbacks and responses and are not used to merely help in validation. Use of prototypes for such purposes is established in innovation literature and many authors believe that as users/customers often cannot articulate or know what they desire beyond current experience and practice (Ries, 2011; Cooper et al., 2002), prototypes with the appropriate degree of richness (i.e. the extent of physical representation, scope, refinement and interactivity (Bradshaw, 2010)) would enable significant insights/evidence to emerge.

Bradshaw (2010) argues that “most authors agree that low-fidelity demonstrators/prototypes are more suitable during the early stages when the emphasis is on understanding requirements and specification development, whereas high-fidelity prototypes are thought to be required for more rigorous testing and uncovering usability problems in the later design stages. The exception appears to be designs that include a physical manipulation or interaction where some element of feel is involved.”

5.2.4.2 Iteration

Iteration is also another critical factor that influences the success of developing innovation using data analytics. As Studds (2014) states, organizations that have recognised the value of experimentation and constant iteration, have extended their analytics focus from one-off benefits to continual improvement.
Walmart stresses the importance of launching a product or service early and hence iterating to improve it: “Strategically, we believe that both the idea and execution are valuable, but we believe that the idea needs to be executed and implemented in the real world. That’s how we view innovation in our analytics team, where we consider an idea valuable only if it is ultimately executed. The idea on its own can be a starting point to begin the learning experience and iterate, but even more valuable feedback and learning will be possible when the idea is implemented, and the iteration happens in real world.”

Similarly Google believes in the power of iteration based on feedback, and reports: “For creating innovation using data analytics, our aim is continual innovation as opposed to instant perfection. The best part of working on the web and with analytics is that we are able to conduct experiments continuously and change things rapidly. For example, the first version of AdWords, which was released in 1999, was not successful but the reason why that eventually became one of our best products is that we continuously iterated and ultimately reached the model that is operating today. We are still iterating on the current model and we run tens of thousands of search and ads quality experiments every year and launch new products based on the results.”

Amazon reports: “Our iterative process frequently enables us to learn invaluable lessons. Being able to observe users use our products in their natural settings, is the best way to determine what works and subsequently act on the feedback. It is always our goal to learn early and effectively about our products and the user behaviour around them, because we can utilise these learnings and respond to them as opposed to travelling down a wrong path for too long.”

Similarly, Airbnb reports: “Iteration is at the core of everything we conduct at Airbnb and it is also crucial when using data analytics for innovation. We were not the first to the market, however, we were able to progress significantly in the market by acting extremely quickly and learning promptly through rapid iterations and grounding our decisions in data.”

Taking an iterative approach to innovation is suggested by many authors, where in order to create, deliver and validate the user/customer benefits (i.e. the value proposition), the development is conducted in short-time increments in order to enable the market and business assumptions to be constantly tested and alterations or refinements to be made nimbly (i.e. experimentation and development are done concurrently). Bradshaw (2010) further adds to this view and explains that “launching early and iterating also helps to build new capability by providing a mechanism
Field studies: case study analysis
to simultaneously test organisational capability, market assumptions and technical feasibility, thereby allowing a new form of concurrent learning to develop”.

The goal is to gradually develop and learn via deliverables, capabilities and prototypes that become increasingly richer (Liedtka and Ogilvie, 2011) and therefore, create a product/service based on facts about the desirability, feasibility and viability of a solution (IDEO.org, 2015) rather than executing a marketing and sales plan (Blank, 2013). This is in order to rapidly prove the existence of a market that reacts positively to a certain solution’s value (i.e. a set of paying customers or frequent users), before committing and scaling resources prematurely in the wrong route to market (i.e. wasting them). Therefore, the purpose is to efficiently achieve ‘product-market fit’, or in other words, evidence that the product and service offerings are indeed creating customer value and are obtaining traction in the market (Osterwalder et al., 2014; Torres-Padilla et al., 2015).

5.2.4.3 Fast-paced execution

Fast-paced execution is another influencing factor derived from our case studies, where evidence of fit is obtained by using physical 3D or 4D/live (real time) prototypes or minimum viable products that continuously increase in fidelity (e.g. that allow to experience the functional and emotional features). The aim is to learn about the value of a solution for potential users/customers as well as other stakeholders such as suppliers and employees, by accelerating the transformation activities that lead to value creation for customers, and responding (as best as possible) to questions such as those suggested by Torres-Padilla et al. (2015): what is feasible to be developed? What attributes or features of the solution are valued most by the user/customer (if any)? How positive do they rate their experience with the solution? Are they returning? Are they willing to pay in any revenue scheme? What should be sourced internally and what externally? What are the potential suppliers/partners, capable of delivering, in terms of the required quality and scale? Are they prepared to partner under certain financial (and other) conditions? What is viable that fits within the organisation’s strategy? What should be the marketing and sales strategy, process and resources for a successful execution? In summary, what should be the business model over time that fulfils the stakeholders expectations?

Wells Fargo reports: “Twice a year our team in analytics, uses ethnography studies, customer councils and insights collected from internal metrics to identify ten priorities. These priorities are subsequently used to develop, test and deploy new products and services every three month. Our industry is changing at such a rapid pace, that it is essential for us to move fast and iterate through products and ser-
vices to perfect them, as opposed to try to establish perfection before a product or service is in the market. Our non-traditional competitors are moving extremely quick, innovating and delivering innovative products and services in exceptionally short cycles, therefore we are required to be fast if we are to stay competitive and iteration (and learning from it) is one way to accomplish that."

Amazon further emphasises the importance of fast-paced executions, and reports: “The reason why we are able to innovate rapidly is a concept referred to as “hill climbing”, which allows every single person in the organisation, regardless of seniority, to be able to access data and tools to test their ideas and intuitions. The importance of data in decision making is obviously established, however, combining that with the ability to test early and frequently, enables rapid innovation and tremendous value creation becomes possible at scale. “. The concept of employees being able to test their ideas and intuitions, is also frequently used in many other organisations such as Facebook and Google.

Moreover, Facebook reports: “We always encourage entrepreneurial innovation, where Startups can move fast and take risks of kinds and magnitudes that their larger counterparts are unable or unwilling to take. If launching a new product takes months or years instead of weeks, it will be exceedingly difficult to be entrepreneurial and therefore to foster innovation, we have a “launch early and iterate” philosophy and one of the company’s rules of thumb is: “If you are not embarrassed by your first launch, you have not launched early enough.”

5.2.5 Managerial support

The case results show that managerial support is a critical factor for the success of developing innovation through data analytics. This result is consistent with literature reviewed in section 3.4.1, where it was revealed that the role of senior and middle managers was crucial for the success of incumbents in developing discontinuous innovations.

The leaders’ support and guidance is vital in promoting innovative efforts, particularly at the initial creative stage, because it contributes to effective interactions among group members (West et al., 2003), and they are also able to create conditions for the subsequent implementation of innovation (Mumford and Licuanan, 2004).

Both individual and group level characteristics of leaders affect the success of creating innovation using data analytics. At the individual level, important factors that allow for the development of discontinuous innovation include tolerance of am-
Field studies: case study analysis

bigness (Barron and Harrington, 1981; Patterson, 1999), self-confidence (Barron and Harrington, 1981), openness to experience (George and Zhou, 2001; Patterson, 1999; West, 1987), unconventionality (Frese et al., 1999; West and Wallace, 1991), originality (Patterson, 1999; West and Wallace, 1991), rule governess (Frese et al., 1999; Simonton, 1991), authoritarianism (Simonton, 1991), independence (Patterson, 1999; West, 1987), proactivity (Seibert et al., 2001), intrinsic (versus extrinsic) attribution bias (Frese et al., 1999; West, 1987), determination to succeed (Amabile, 1983), personal initiative (Frese and Zapf, 1994), and managerial tolerance of change (Damanpour, 1991).

However, upper echelon theory proposes that composition and characteristics of the top management team generates a stronger explanation of organizational outcomes than a leader’s characteristics alone. The role of middle managers is also critical for developing innovation through data analytics, because most strategic proposals take their fundamental shape at the lower levels of hierarchical organizations.

The categories that emerged from the analysis are managers as team members and adequate resources.

5.2.5.1 Managers as team members

From our case studies it was derived that an important factor for the success of data analytics endeavours is for the managers to not allow their own presumptions or goals to become a barrier in the project. Amazon reports: “without the support of the CEO and executive team and their willingness to challenge their own assumptions, a data-driven culture is meaningless. A top-down approach to this is the most effective way of breaking the cycle of the “HiPPOs” (Highest Paid Person’s Opinions), and instead move through a culture where decisions are made through data, experimentation and iteration. Managers should be”. "HiPPos" mentality breaks the momentum of teams and does not allow for creativity and innovation. It frequently results in team members losing motivation and can be destructive for the dynamics of the team.

Furthermore, managers should not be regarded as judges who only make decisions on projects in isolation and instead should be taking part in projects as team members and contribute to creativity generation. Facebook reports: “At Facebook, top executives including CEO Mark Zuckerberg are involved as entrepreneurial thinkers as opposed to judges and essentially our innovation process is about having these thinkers participate like they are other people on the team instead of only providing approval/disapproval. The reasoning for this is that it is difficult to judge something
if you are not part of the process of creating it and if leaders are involved in design decision-making then decisions don’t come as a surprise.”

5.2.5.2 Adequate resources

Allocation of adequate resources to analytics projects is an important factor affecting the success or failure of developing innovation. In 3.4.2.2 the review of innovation literature indicated that resource allocation is particularly important for creating discontinuous and disruptive innovations. Visa reports: “Management awareness and understanding of the business potential of big data and analytics are especially critical for the success of analytics projects. Hardware, software, and programming or analytics skills are all becoming increasingly available due to the large amount of educational material and degrees available. However, many managers still do not appreciate the true value and potential of big (and small) data, which is problematic because senior management sponsorship is necessary to proceed aggressively with this resource, and therefore, educating managers is one of the most important thing that organizations can do.”

The influencing factors that affect the resource allocation in organisations, typically include absolute and relative R&D intensity (Parthasarthy and Hammond, 2002), commitment to differentiated funding (White, 2002), annual turnover of resources (Mohr, 1969), and slack resources (Damanpour, 1991; Kanter, 1983; O’Brien, 2003).

The main inhibitor for resource allocation to discontinuous innovations such as those potentially possible through data analytics, is structured routines (Nelson and Winter 1982) like the key evaluation factor of financial returns (Christensen 2006) and traditional market research reports. These structured routines are particularly difficult to alter once established and constrain resource allocation to data analytics projects as they apply the same criteria for them as existing businesses. To solve such resource allocation problems for discontinuous innovations in general, some scholars have conducted empirical research that have proposed several solutions, such as using strategic buckets to manage sustaining versus disruptive projects independently (Chao and Kavadias 2007; Hogan 2005) or having projects in all the various pipeline phases and managing each phase as a mini-project (Hogan 2005).

5.2.6 Use variety of data

Using a variety of data sources in analytics projects, was also indicated as a significant factor contributing to the success of innovation through analytics.
Wells Fargo reports: “The explosion of Fintech companies has occurred over the last few years and they are utilising data in extremely creative ways around customer behaviour, especially by combining various non-traditional data such as social platform data. We have also been doing this, however mostly to enhance our core offerings (such as fraud detection), and not to reshape the business. This competitiveness of the industry means there is a push for the banks to innovate and create and use more data for innovation. It will create new value for the customers and I think it is an enormous opportunity for all of us.”

Visa reports: “To use analytics to its full potential and enable innovation, companies must use a variety of data in analytics processes, especially internal data that typically exist in silos. Siloed data can limit how effectively information is analysed across an organisation, because risk data is generally not accessible from marketing and product information, which limits the depth of information that can be obtained regardless of the available volume.”

Walmart: “For us the main benefit of the use of big data for innovation is that we have access to not only structured data but various types of unstructured data, such as texts, images, and voice recordings, and in addition we can determine correlations within the data. This knowledge of complex processes and interrelationships, has the potential to create significant customer and society value and therefore the objective is to obtain patterns and to create accurate digital depictions of the real world. To enable computers to understand language, you also need significant amounts of data and in-depth knowledge of the world. Using a variety of data is where real value is created.”

In a study conducted by IBM in 2013 on 11 percent of the global financial industry respondent pool, more than four out of five banking and financial markets respondents with active big data efforts are analyzing transactions and log data that are machine-generated data produced to record the details of every operational transaction and automated function performed within the bank’s business or information systems – data that has outgrown the ability to be stored and analyzed by many traditional systems. Consequently, in many cases this data has been collected for years, but not analyzed.

Where banks and financial industry firms are behind their cross-industry peers is in utilizing more varied data types within their analytics pilots and implementations. According to the IBM study (2013) slightly more than one in five (21 percent) of these firms is analyzing audio data (often produced in abundance in retail banks’ call centers) while slightly more than one in four (27 percent) report analyzing social data
Field studies: case study analysis

(compared to 38 percent and 43 percent, respectively, of their cross-industry peers). Most industry experts believe this lack of focus on unstructured data is associated with the challenge of integrating the organizations’ structured data. In order to obtain the maximum value from analytics, there is a need to analyze multiple data types. Examples of this include using advanced capabilities designed to analyze text in its natural state, such as the transcripts of call center conversations that include the ability to interpret and understand the nuances of language, such as sentiment, slang and intentions, and are often used to bolster efforts to understand behavior and preferences and improve the overall customer experience. Even more complex types of unstructured data, including geo-spatial location data, voice data, video or streaming data are also important sources of insights.

The ability to achieve significant and measurable business value from big data can only be realized if organizations put into place an information foundation that supports the rapidly growing volume, variety and velocity of data (IBM, 2013). Walmart, for example, generates more than a million customer transactions each hour that are imported into databases estimated to contain more than 2.5 petabytes of data. The company is capable of combining data from a variety of sources including customers’ past purchases and their mobile phone location data, Walmart internal stock control records, social media and information from external sources such as the weather, and initiate tailored sales promotions.

Walmart reports: “Every project typically begins with a goal and with objectives to reach that goal and data analytics projects are no different. However, defining the goal can be a difficult process, especially when the goal is imprecise and is little more than the statement “using the data better”. It is essential to define the goal and the objectives i.e. the exact means by which to reach the goal, before hunting for data sources because it will be necessary to gather information from a multitude of sources, some internal and others external that may have to be purchased or may be available under the public domain.”

5.2.7 Organisational structure

Organisational structure includes the categories affecting the structure of the organization that could be adjusted or changed to suit the changing environment. Two categories that emerged from the analysis include the physical environment of the firm and the existence of separate small business units dedicated to innovation.
5.2.7.1 Physical environment

The role of the physical environment in promoting creativity and collaboration was emphasised in our case studies. In fact, the majority of the firms that participated in our study have gone to extreme lengths to preserve the vibrant and collaborative atmosphere in their transition from start up through to incumbent. The effect of physical environment on creativity and innovation is well-established in literature. As Webber (1993) states, providing individuals and teams with an energising and dedicated innovation space to focus, think, act, share, reflect and decide throughout the projects, facilitates the development of the required strategic conversations and trust. It also enables people to distance themselves from the paradigms, rules and routines of operations in the exploitation arena, and move towards exploratory behaviours (Hendry and Seidl, 2003; Lewis and Moultrie, 2005). The layout of the space where certain activities are conducted, can also be configured in a way to promote the required behaviours and interactions (Doorley and Witthoft, 2012).

Airbnb reports: “To stimulate creativity and allow our designers to put themselves in the shoes of the user, our entire building is purpose-built. We have spaces that are designed to replicate a room that is located in France, or Vietnam, has a river outside, or is next to a waterfall. We also have pictures of our customers and their unique stories across all our walls. These settings assist with promoting creativity and empathy tremendously.”

Facebook reports: “The physical environment has a significant impact on how people think and feel. At Facebook, engineering, management, and other teams frequently move their physical desks and furniture in order to cultivate new and innovative ideas by joining new groups — this happens in person, and on a daily basis rather than moving back and forth from permanent desk locations.”

5.2.7.2 Separate small units

Having small units dedicated to developing innovation through data analytics and creating innovative products and services that result in significant business or customer value, is another success factor identified in the case results. Apple reports: “one way that you can increase the probability of succeeding with analytics is to create small, cross-functional innovation teams in each region, that are employed exclusively to search for breakthrough ideas that can add value.”

Wells Fargo similarly believes in the importance of having dedicated and small innovation groups, and reports: “The advantage of a smaller group is that it connects well across an extremely large company like ours and enables the generation of newer
and more innovative ideas. Therefore, for our analytics innovation projects we have a few small teams working collaboratively. One of these teams is a data analytics innovation group inside the bank that aims to be the catalyst for the bank to change its business models and processes. Our aim is to provide these small teams with the resources, investment and support required to enable the occurrence of something transformative.”

Google reports:” Innovation has consistently been driven by a person or a small team that has the opportunity of thinking of a new idea and pursuing it. We believe Innovation occurs when people are not forced or under pressure therefore it is crucial that our people have time for reflection. In our case, we attempt to promote and encourage innovation and true outside the box thinking using principles such as “20 percent time” (that allows Google employees to spend 20% of their time on any project they are interested in), and the small technology teams, which are not directed.”

The case results are consistent with prior research discussed in 3.4.2.3. Innovation research has largely focused on firm or business unit size as a key success factor for R&D effectiveness (Cohen and Klepper, 1996; Tsai and Wang, 2005) and the recent trend argues that R&D investments are more productive for small than for large firms or business units when introducing new products (Lee and Chen, 2009; Lejarraga and Martinez Ros, 2008). In addition, in the research on disruptive innovations, case studies and surveys in high-tech industries have also shown that the size of the firm or business unit is negatively correlated to the success of disruptive innovation (Christensen and Raynor, 2003; DeTienne and Koberg, 2002; Tushman and O’Reilly, 2002).

5.3 Framework for innovative analytics

Using the success factors derived from the grounded theory (section 5.2.), I created a framework for conducting data analytics for the creation of innovative products and services, which is named Creative Data Analytics (CDA) framework. The main objective for the creation of this framework was to achieve a clear and simple mapping of the identified success factors to stages of the creative data analytics process, in a way that each of these elements can be implemented in any financial organisation to increase the probability of their analytics efforts resulting in innovation.

CDA aims to articulate a holistic and end-to-end solution for strategic data analytics innovation management, and therefore its design incorporates ‘best of breed’ features
from established and progressive approaches that have proven their value, and it also facilitates the integration of tools/toolkits that such approaches provide.

In the grounded theory developed in section 5.2., it was deduced that the goal and objectives of the data analytics projects i.e. the problem to be solved, is a fundamental success factor for innovation and value creation. The underlying notion of relating the problem space and the solution space by the identification of a main concept is relatively well established. Dorst and Cross (2001) explicate that creative design is not a simple process of first defining the problem and subsequently searching for a satisfactory solution but rather a process of developing and refining simultaneously both the formulation of a problem and ideas for a solution. This is achieved by continuous iteration of analysis, synthesis and evaluation activities between the problem space and solution space. The authors use a ‘co-evolution’ model proposed by Maher et al. (1996), to describe this pattern of development in their experiments: “A rough description of what happened in this case is that a chunk, a seed, of coherent information was formed in the assignment information, and helped to crystallise a core solution idea. This core solution idea changed the designer’s view of the problem. We then observed designers redefining the problem, and checking whether this fits in with earlier solution-ideas. Then they modified the fledgling-solution they had.” This co-evolution model is a fundamental principle in CDA to ensure that it assists organisations when concepts and problems are not defined, which is mostly the case when using analytics for innovation.

It was also deduced from the grounded theory that understanding customer needs and creating products and services in response to customers and users’ latent needs will increase the probability of success of data analytics project. Learning through iteration and prototyping, experimentation and fast-paced execution were also identified as critical success factors for data analytics projects creating significant value. As explained previously, my aim is to use “best of breed” approaches to incorporate the success factors derived from the grounded theory, in the CDA framework. One methodology that combines human centred empathy (creating products and services based on users’ needs), prototyping and iterative learning, is Design Thinking. In section 3.4.2. of this thesis, the methodology of Design Thinking, which is a widely popular human-centred approach to innovation, was discussed. To incorporate the above mentioned key success factors and enable user-centric solutions, we integrate Design Thinking in the CDA framework.

Furthermore, lean innovation approach is adopted and strengthened by the compatible agile and experimentation philosophies, techniques and tools, in order to increase the efficiency and effectiveness of CDA. Team focus and engagement, in
addition to continuous user/customer feedback and facts are necessary for avoiding waste and creating value fast (Sehested and Sonnenberg 2011; Ries, 2011). An agile development philosophy, which was developed by the software industry, emphasises the significance of individuals and interactions, working products, customer collaboration, and responding to change rapidly and nimbly (Beck et al., 2001; Satpathy, 2013). It utilises short time-boxed increments in which a deliverable is continuously demonstrated to stakeholders (rather than documentation) (Cooper, 2014), which fits well with the experimentation approach that aims to incrementally validate with stakeholders the primary business assumptions that are the basis of an envisioned solution concept.

The CDA framework has been created based on two fundamental aspects of conducting data analytics for innovation, which were derived from our research findings:

1. Prerequisites: covers all the activities that are to be placed prior to the start of the analytics project, in order to increase the probability of success of the creative analytics process. These activities include the derived success factors of: top management support and commitment, innovation strategy and innovation culture, teams, and the organisational structure.

2. User-centric creativity combined with analytics: Understanding business and customer needs and pain points, prototyping, testing, iterations and learning by doing, are factors that significantly affect the probability of achieving innovation through analytics. This part of the process model brings together both creativity and analytics techniques while ensuring an iterative model that enables learning through experimentation, prototyping and testing is in place. The main aim of this part of the CDA framework is to conduct data analytics effectively. Therefore, in this phase of the process model, the data analytics process model discussed in chapter 2 is also used as the baseline. This is primarily because CRISP DM framework is the most widely used data analytics process model in the industry.

These two main elements correspond to the stages of the Creative Data Analytics framework. Each element has its specific activities and phases that are devised according to the success factors derived from our research and by building on the established and progressive approaches of Design Thinking (e.g. Liedtka and Ogilvie, 2013; IDEO.org, 2015), Agile Development (e.g. Satpathy, 2013) and Lean Startup (Ries, 2011; Maurya, 2012; Blank, 2013); and also utilising the CRISP DM Data analytics life cycle, which was chosen as it is currently the most frequently used and most popular data analytics process.
Adopting these approaches enables the delivery of their benefits that address various management needs and success factors for innovating through data analytics - by combining them in a single and coherent framework that allows for the purposeful and simple integration of tools along an innovation cycle. In addition, CDA provides a means of applying unique configurations of tools that enable activities to be conducted in a series or parallel fashion as required, and allow evaluation points to be flexibly configured and integrated with strategic portfolio management. The Creative Data Analytics framework created is illustrated in figure 5.1.

It should be noted that the activity flow in the CDA framework is a simplification of a complex phenomenon where activities often ensue in a non-linear and iterative fashion, creating new opportunities and ideas (i.e. new ‘triggering events’), which act as suitable starting points for new innovation initiatives. Furthermore, the aim of CDA is not to display all the possible activity interconnections and iterations that
might occur in data analytics projects but to provide a simple basis to support the
design of ‘project process expeditions’ (e.g. tool selection, configuration and inte-
gration) within the logic of a CDA project. The arrows indicate the most important
and frequent dependencies between phases.

CDA is not intended to be a linear process. Such process would be undesirable
for our intent, which is discovering and devising problems that are undefined or
poorly defined at the beginning of the projects. In fact, the CDA framework is
most appropriately applied in situations in which the problem, or opportunity,
is not well defined, and/or a radical idea or concept is required, that is, an idea
that has a substantial and positive impact, such as creating a new market or en-
abling major revenue growth. Therefore, instead of following a linear process, the
aim of the CDA approach is to define the problem and create potential solutions
promptly—considering that our knowledge is imperfect and that these solutions will
be incomplete and potentially flawed—and subsequently use these initial solutions
as a means of facilitating learning and developing more refined insights that can
ultimately result in creating better solutions. This iteration is because the CDA
framework is created based on the belief that an innovation process is always to a
certain degree a journey of discovery with new knowledge being acquired through-
out the project. Therefore, the framework and associated tools have the purpose of
assisting practitioners in plotting their course through an uncertain terrain, while
knowing that the path and the destination itself may change as the route unfolds
(Goffin and Mitchell, 2010). CDA assists in avoiding the trap of investing a sig-
nificant number of resources too early in a project toward the development of a
specific, single solution, and instead encourages several “little bets” (Sims, 2013)
about customer insights and possible solutions. Sims (2013) defines these little bets
as “low risk actions taken to discover, develop, and test an idea.” These little bets
increase the probability that a project team will converge on solution concepts with
the highest potential market success fast.

Consequently, CDA is best described as an iterative approach to problem solving
using data analytics, rather than as a sequence of steps. The specific number of
iterations depends on each project and is, to some extent, unknown at the beginning
of a project and is later established based on the objectives and constraints of the
project, as well as the perceived progress of the work. One of the main undertakings
of the team and its leader throughout a project, is to make a decision in regards to
the best ways to proceed with the project.

CDA has been designed to be modular, scalable and applicable at an appropriate
level of granularity/detail, in order to enable an innovation project process to be
Field studies: case study analysis

described and managed from macro to micro perspectives, while supported by management tools/toolkits (Kerr et al., 2013). From a macro perspective, CDA can demonstrate how the general process will be conducted, by presenting the broad steps that an organisation is required to take. As discussed earlier, an iterative and adaptive approach is suggested. This is in order to improve the management of uncertainty/risk and investment prioritisation, by providing evidence obtained from building, testing and learning through experiments (e.g. demonstrators/prototypes).

Although, loop-back mechanisms such as iterations and pivots, may initially appear to delay the stages of an innovation endeavour, especially the front end of innovation, several authors claim that they assist in preventing waste (e.g. Ries, 2011; Maurya, 2012; Blank, 2013), and fast-track innovation introduction by reducing the total cycle time (e.g. Koen et al., 2002; Cooper, 2014), provided that the projects are adequately resourced and activities are flexibly configured throughout the stages (Cooper, 2014). From a micro perspective, CDA deals with a level of detail associated with the tools, techniques and procedures that are supplied into the macro level to achieve the goals of the organisation.

Flexibility, iteration, experimentation and fast execution are the basis of CDA principles. Even though generally the term efficiency refers to making better use of resources (e.g. funding, materials and time) and the term effectiveness refers to the ability to accomplish a desired result, it should be recognised that their meanings vary between an innovation perspective and that of production/operations. The mutual aspect of both perspectives is the desire to eliminate waste – defined as all activities and tasks that do not contribute to value creation- which is a fundamental principle of the lean approach that was originated in Toyota corporation which resulted in improved effectiveness and efficiency not only in production environments but also in innovation (Ward et al., 1995; Sobek et al., 1999). Since then, Lean approaches have been adopted by many firms in order to improve their innovation processes.

Arguably the most critical distinction between lean innovation and lean production is that innovation is about ‘learning’ and operations are about ‘execution’. In innovation, activities that contribute to value creation include tasks that determine who the users/customers are, what is valuable to them, and how to produce and deliver that value; as opposed to operations, which are concerned with activities that optimise and ensure the delivery of actual value to the customers; and anything else is considered waste (Ries, 2011).

Since the focus of CDA is on problems or opportunities that are not well-defined, lean innovation is appropriate, however, when a valuable solution has emerged and
has been validated by a set of customers, the operational processes can become the priority in order to reproduce it and ensure its consistency. In routine operations, by definition, variation is non-value creating (i.e. waste) and therefore the solution must be reproduced as accurately as possible. However, this is in contrast with an innovation process where the end result is unknown to a certain extent, and therefore flexibility is critical. Thus, in innovation there is both value-creating and non-value creating variation: certain variation is valuable as an essential prerequisite for creating something novel, however, non-value creating variation such as making collaboration more difficult, creating misunderstandings and unproductive confusion represent waste and should be eliminated (Sehested and Sonnenberg, 2011).

In the following sections each of the phases of the CDA framework and their associated activities will be described.

5.3.1 The prerequisite phase

5.3.1.1 Managerial support

As derived from the grounded theory, managerial support is crucial to the success of data analytics projects. In this subsystem, it is important to thoroughly understand the values, interests, motivations and emotions of managers and all other stakeholders involved, in order to ensure that managerial support is maintained throughout the project and adequate resources are available.

One way to gain managerial support and involvement from the beginning of a creative data analytics project is to utilise a brief, which is the classical starting point for the inspiration phase. The brief is to be created with the managers (and other stakeholders) and consists of several mental constraints that provide the project team with a framework from which to begin, benchmarks for measuring progress, and a set of objectives to achieve—such as price point, available technology, and market segment (Torres-Padilla et al., 2015). However, in the same way that a hypothesis is not equivalent to an algorithm, the brief does not serve the purpose of providing a set of instructions or attempting to find a solution to the question before it has been posed. Instead, a well-constructed brief enables flexibility, unpredictability, and experimentation—the creative realm that enables breakthrough ideas to emerge. The project brief is critical in providing a common ground and understanding, between the managers and teams, and it requires a certain level of granularity; if the brief is too abstract, it risks leaving the project team wandering, and if it possesses constraints that are too narrow, it is almost guaranteed that the outcome will be incremental and, likely mediocre (Torres-Padilla et al., 2015).
As Liedtka and Ogilvie (2011) state, the concept/project brief facilitates the integration of vital knowledge that is required to be considered to support the management and communication of a formal innovation project. The concept/project brief is envisioned to be an evolving device; therefore, its content will be enriched, altered and/or adjusted by managers and team members as progress is made, and will act as a continual connector between the managers and the project team.

As with every tool in CDA, the configuration of the brief is flexible and depends on requirements and context of the specific project. The format of the concept/project brief varies for each project, however, some common elements are provided below as examples (Torres-Padilla et al., 2015):

- Opportunity Vision / Scenario / Description
- What is it about?
  - Purpose / Goal / Intent / Scope
  - Target Users / Potential Market Requirements / Constraints / Design Criteria
- Why is it significant?
  - Rationale / Justification (e.g. Need, Quantified Value, Potential Benefits)
  - Potential competitors
- How can it be conducted?
  - Resources (e.g. People, Time, Money)
- When should it be conducted?
  - Plan / Actions / Next Steps Milestones / Deadlines Expected Outcomes Success Metrics
- Who is interested? Who is responsible? Stakeholders / Ownership
- Key Assumptions and Risks

Therefore, the brief comprises of a prioritised list of the assumptions that are required to be true in order for an opportunity concept to work (i.e. validated desirability, feasibility and viability) and the experiments that could be conducted to validate each assumption, including success metrics and the resources required (Ries, 2011; Liedtka and Ogilvie, 2011; Osterwalder et al., 2014).

Analysing the outcomes related to an assumption enables the learning required to pursue one of the following routes: a) conduct further experiments if uncertainty
remains; b) validate the assumption in order to support further progress in the planned direction; c) invalidate the assumption and pursue a ‘pivot’ (Blank, 2013; Ries, 2011) that is an iteration in search of the right fit in one/all of the stages of the data analytics project (e.g. problem-solution fit, product-market fit or business model fit) (Osterwalder et al., 2014). Any major alteration in the logic of the business concept (e.g. a different customer segment) would be reflected in the roadmap and project brief, and may lead to the testing of new assumptions.

The main aim of this pre-requisite subsystem, is to provide the managers with a clear understanding of the project as well as enabling the team members to gain adequate knowledge of the managers and their desires in order to enable the project to be conducted smoothly and with the managers’ support. The purpose of this subsystem is also to learn about the enterprise and the customers and it facilitates exploration activity leading to new insights, responding to questions such as the following:

- **External environment (intelligence):** what are the key trends affecting the organisation (e.g. STEEPLE)? What are the forces governing competition (e.g. Porter analysis)? How is the market and what is it demanding (e.g. user/customer and market research)? What technologies might be of particular interest?

- **Future (foresight):** how the environment might be in the short, medium and long term (as predicted from current reality)? What are the central certainties and uncertainties about the future that are relevant to the organisation?

- **State of the organisation:** what are the current capabilities and resources of the organisation? What is its business model (i.e. how it creates, delivers and captures value)?

To facilitate the discovery of answers to these questions, during this subsystem various key functions are performed, that include gathering and capturing pieces of data and information, as well as foreseeing potential negative aspects (e.g. risks, difficulties and/or assumptions) and/or positive aspects (e.g. benefits and value) associated with pieces of information; and subsequently making sense of all data/information by clustering it into meaningful insights that can be assessed and prioritised (Torres-Padilla et al., 2015).

In practice, this subsystem can be executed with the aid of tools and techniques such as the following:
Field studies: case study analysis

- STEEPLE (e.g. integrated within Roadmap) (More et al., 2015) in order to discover social, technological, economical, ethical, political, legal and environmental trends and drivers.

- Scenario planning (e.g. Ringland, 2006) to identify potential futures by analysing today’s signals in the environment.

- Porter’s forces (competitive analysis) (Porter, 1979, 2008) for understanding the level of competition and factors involved within an industry that may affect the ability of an organisation to serve a market and make a profit.

- LEGO Serious Play (e.g. Frick et al. 2013) to understand values, interests, motivations and emotions of stakeholders (even unconscious or hidden drivers) and their roles, relationships and potential impacts within a team, environment or system.

- Business model canvas (Osterwalder and Pigneur, 2010) to understand a current business model through its main elements and interrelationships.

5.3.1.2 Innovation culture

Culture was also identified in the grounded theory as a critical success factor for enabling innovation through data analytics. The culture component of the CDA framework is dedicated to the overall pattern of behaviour of people in the data analytics innovation project as part of a larger organisation system and includes aspects such as shared symbols, habits, norms, values, beliefs and assumptions that enable individuals in the system to interpret and act upon their environment (Torres-Padilla et al., 2015). Therefore, it is vital to comprehend how a culture has been formed and how it could be altered if strategic goals are to be achieved (Schein, 1984). However, culture typically reflects the imprint of earlier periods in a persistent way, and therefore change may not be simple (Marquis and Tilsik, 2013). Furthermore, the ideal ‘ambidextrous’ character of a firm necessitates at least two broadly different ‘sub-cultures’ that are required to be dealt with, which are innovative culture and operational culture. While an operational culture often values characteristics such as precision, efficiency, low risk, control and quality, an innovation culture – which is required for creative data analytics projects - values characteristics such as flexibility, novelty, risk taking, agility and experimentation.

Creative data analytics initiatives require a few primary conditions to be in place including an explorative culture, creative people, and autonomous (self- managed) structures/teams to facilitate the ability to go beyond current paradigms and mind-sets. Consequently, an understanding of the objectives and current level of maturity
Field studies: case study analysis

of various capabilities enable making decisions on what enabling elements to develop and to what extent, as well as when they should be implemented. This is a challenge of varying level of difficulty depending on the type of organisation, for example, for start-ups, a culture of creativity, experimentation and risk-taking is typically more common and natural than in established large organisations, where a culture of precision and risk-avoidance, driven by short-term objectives, is more common (Torres-Padilla et al., 2015).

Furthermore, one of the major impediments to adopting an exploratory culture is fear of failure. For many organisations, it is difficult to accept the concept that experimentation and/or failure are vital for the long run success of the business, and failures should happen early and act as a source of learning. Creative data analytics requires a strong innovation culture with experimentation at its heart, that also encourages prototyping (quickly and inexpensively) as part of the creative process and not just as a means of validating finished ideas.

5.3.1.3 People aspect

The people component of the CDA framework is dedicated to the emotional, social and intellectual ‘assets’ of people that include motivation, style, knowledge, skills and experience. The interdependence of these assets, to a large degree determines the character of an innovation system, and an adequate mix and interconnection of people permits teams to think and conduct the ‘right thing’ quickly in novel situations (Weick and Roberts, 1993). Consequently, the CDA framework should be reinforced by human resource processes to recruit, train, monitor and reward people in alignment with the innovation strategy (Goffin and Mitchell, 2010; Dornberger and Suvelza, 2012). Based on the particularities of each project, dedicated people with the right profile and in adequate numbers should be allocated to lead and/or perform roles in structures and processes. In contrast to a business operation that is often associated with authoritative managers with an operational excellence track that rewards people for the achievement of margins and productivity (e.g. predominantly measured with quantitative targets), an innovation system is required to select visionary leaders with an entrepreneurial spirit, and reward people for learning from failures and for achieving the milestones and growth (e.g. measured with both qualitative and quantitative targets), particularly when more radical approaches to innovation are desired (Torres-Padilla et al., 2015).

Multidisciplinary team structures of an appropriate size and combination of authority, drive, knowledge and skills are typically used to drive innovation, and are also important for driving innovation using data analytics. Furthermore, to enable di-
vergent thinking, it is critical to have a diverse group of people—with the capacity and the disposition for collaboration across disciplines—involved in the CDA process (Cohen and Bailey, 1997).

To operate within the interdisciplinary environment of a CDA process, an individual is required to possess strengths in two dimensions (the “T-shaped” person): 1) a depth of skill that enables tangible contributions to the outcome. 2) empathy for people and disciplines beyond one’s own, which is often expressed as openness, curiosity, optimism, a tendency toward learning through doing, and experimentation (Brown and Wyatt, 2010).

The above-mentioned structures for teams are associated with faster development rates (Stalk and Hout, 1990; Eisenhardt and Tabrizi, 1995) and enhanced performance (Brown and Eisenhardt, 1995). The use of these types of teams has also been connected to enhancements to organisational efficiency and quality (Applebaum and Batt, 1994), and financial performance (Macy and Izumi, 1993). In addition, they enable the flexibility required to effectively and efficiently respond to the continuously changing demands of the organisation’s environment (Zaccaro et al., 2001) and facilitate the development of the various human dimensions involved during the strategy-related activities that include cognitive, emotional and social (Roos and Victor, 1999; Kerr et al., 2012).

5.3.1.4 Physical environment

Dedicating a special space to the CDA projects influences the perspective of people and places the shared goals front and centre. By providing individuals and teams with an energising space to focus, think, act, share, reflect and decide throughout the whole innovation cycle of the CDA process, the dedicated space facilitates the development of the required strategic conversations and trust (Webber, 1993). Furthermore, it alters people’s thinking and enables them to move away from the paradigms, rules and routines of operations in the exploitation arena, towards exploratory behaviours (Hendry and Seidl, 2003; Lewis and Moultrie, 2005). The layout of the space where certain activities are conducted, is configured to promote the required behaviours and interactions (Doorley and Witthoft, 2012). The dedicated space also constantly reminds senior management of the significance of innovation investment for survival and growth. We refer to this dedicated physical space for CDA as “Innovation Room”.

Most of the management tools, that provide the primary means of practical implementation (Shehabuddeen et al., 2000), are brought to life and utilised in the
“Innovation Room” where participants configure and integrate the tools and related techniques to facilitate tangible outcomes along the CDA innovation cycle. Because tools are the fundamental foundation for results, their development must consider comprehensive principles such as those set out by Kerr et al. (2013) and be: 1) ‘human-centric’ so that all team members are able to participate completely; 2) activity-based and utilise workshops and small-group activities with structured tasks; 3) if the activities are facilitated, the facilitators must be neutral and their focus should be on the process as opposed to content; 4) as ‘lightweight’ as feasible in order to facilitate a degree of flexibility and avoid being excessively prescriptive; 5) modular to facilitate their simple integration with one another; 6) scalable so that they can be used at different hierarchical system levels both inside and outside the organisation; and finally 7) visual, both when they are being used (e.g. workshop charts and templates) and when communicating the outputs (e.g. reports or summaries).

The use of resources, materials and knowledge – pictures, posters, data, ideas, workflows, and prototypes – that are related to a project on display in this shared space, results in the project participants to be continuously stimulated by them. Similarly, utilising flexible furniture such as sofas, chairs, tables on wheels, and moving whiteboards can also assist. However, instead of just applying these two criteria without much consideration (e.g. simply using wheeled furniture in all rooms), it is important to think thoroughly about what kinds of behaviours one intends to encourage with a given configuration, because the configuration sends a message that tells people how to behave, even if people are frequently unaware of it (Doorley and Witthoft, 2012; Britos Cavagnaro, 2013; Torres-Padilla et al., 2015).

The fundamental concept of visual management, which is the basis of the ‘Innovation Room’, is not new and has been regularly implemented under different names and for various specific purposes, the examples of which include: the ‘Chart Rooms’ of the early twentieth century (e.g. Yates, 1985), the ‘War Rooms’ of the second world war, and the more recent ‘Obeya Rooms’ (e.g. Warner, 2002), ‘Pulse Rooms’ (e.g. Kaya, 2012), ‘Design Studios’ (e.g. Britos-Cavagnaro, 2013), ‘Creative Spaces’ (e.g. Doorley and Witthoft, 2012), and ‘Innovation Labs’ (e.g. Stevens and Moultrie, 2011). The primary and common purpose of all these examples is sharing knowledge and experiences, in a focused, agile and visual way in order to improve decision making and outcomes.

The value of the principles deduced from these experiences is well recognised and has been integrated in CDA framework to assist in accelerating the innovation development using data analytics. However, it should be noted that this approach is not
free of challenges and limitations. A low-tech approach specially for a data analytics project, is occasionally criticised as childish and weak – which represent misunderstood perceptions because the process is entirely serious, where problems are openly discussed, and decisions and responsibilities are exposed to everyone (e.g. Sehested and Sonnenberg, 2011). Regarding the critique of being weak, there are a few properties that have been criticised as disadvantages of using a low-tech approach which include (Kaya, 2012):

- Because they are physical, for organisations with several offices it is a challenge to share information with managers that are in various locations.

- As the complexity of the solutions (e.g. products) that are being developed increases, the volume of information that must be included and managed also grows Several challenges are associated with this:

  – Because the information is operated manually, detailed information must also be tracked down manually, which may pose challenges in specific cases as well as being time consuming.

  – The limited space on boards, walls and rooms, do not accommodate for considerable expansions and adding an entirely new project without deleting another may present challenges.

Several solutions that utilise low and high technology have been proposed (e.g. video conference and software applications, digital visual management technology), which can replace or complement the above approach, resulting in hybrid approaches. Generally, these hybrid approaches aim to overcome apparent challenges and issues of physical settings such as: collaboration among multi-site teams, surface cost of physical spaces, and security and reuse of knowledge. But, it is essential to consider to what degree this would improve the total benefit without undermining critical cognitive, social and emotional benefits that come with face-to-face, high-touch approaches (Torres-Padilla et al., 2015).

In general, a physical space that is ‘High- Touch’ (e.g. highly collaborative) and ‘Low-Tech’ (e.g. predominantly manual, supported by wall-charts and sticky notes) is promoted in lean approaches to reinforce closeness and commitment (Sehested and Sonnenberg, 2011).
5.3.2 User-centric data analytics phase

The core activities and deliverables in each stage of the CDA framework are being dynamically defined based on the research of critical unknowns and the validation of essential assumptions, and therefore the framework is more associated with progressive approaches (e.g. Liedtka and Ogilvie, 2011; Maurya, 2012; Blank, 2013; Osterwalder et al., 2014), which are intended to address higher-risk undertakings (such as radical innovation projects) in extremely uncertain and fast-changing environments (Cooper, 2014). To address these higher risk, less defined endeavours, the CDA framework takes advantage of divergent-convergent thinking and flexible evaluation points, as well as tools and techniques to perform each stage of the CDA framework. In the following sections, these concepts are discussed in detail.

Divergent-convergent

Sequential divergent-convergent phases were originated from early work on creativity (Osborn, 1953; Parnes, 1967) and psychology (Guilford, 1967). Divergent thinking can be defined as creating a multitude of responses to an open-ended task or question in which the outcome is not fully developed by the accessible information. It aims to produce a significant and diverse number of alternative responses that comprise of novel, unanticipated, or uncommon ideas and therefore, it is frequently associated with creativity or generative thinking. Convergent thinking on the other hand, involves identifying the most appropriate answer that is conventional to a well-defined problem or question, by examining voluminous facts, information or ideas, for their logical validity, within a set of rules, guidelines or paradigms and therefore it is typically associated with critical thinking. In the context of the CDA framework, the central concept is to initially create a significant amount of new information pieces, insights or ideas (i.e. divergence) and subsequently examine them and focus on the most appropriate one(s) (i.e. convergence). Iterations are continuously conducted between these two modes of thinking as necessary and consequently the divergence-convergence cycle stimulates exploration and learning.

The divergence-convergence scheme is widely popular among authors and practitioners in the field of creativity as well as related fields such as innovation and entrepreneurship and is used in many methods such as Creative Problem Solving (CPS) that established its practical and explicit use within a structured process and toolkit (e.g. Isaksen and Treffinger, 2004), the Lateral Thinking and Six Thinking Hats techniques that enable the composition of creative processes conformed of both phases (referred to as ‘expansion’ and ‘contraction’) (de Bono, 2005), and design thinking methodology that acknowledges the effectiveness of divergent-convergent
thinking for understanding and structuring an innovation process (e.g. Liedtka and Ogilvie, 2011; IDEO.org, 2015). Therefore, many firms have instrumented their innovation processes to varying extents around divergence–convergence (directly or indirectly), because it appears to be an efficient means of structuring thinking to generate a range and quality of outcomes. As an example, in Toyota’s set-based simultaneous approach for innovation development, implicit divergence allows many variations of concepts to be generated and explicit convergence enables focus towards the most appropriate options (Ward et al., 1995; Sobek et al., 1999).

In the context of a CDA project process, the divergence–convergent concept can be applied with various levels of granularity where a process can be characterised with only limited sets of divergent–convergent cycles, or be considerably detailed and comprise of many sets (for example at the tool level, where team interactions with every tool in the process may be facilitated with a divergent–convergent approach).

**Evaluation points**

In addition to having divergent-convergent thinking implemented in the CDA framework, it is also important to have evaluation points that can be flexibly organised throughout the journey to optimise risk and resource management (Cooper, 2014). Evaluation activities facilitate decision making at two levels of project and system/portfolio. At the individual project level, evaluations may be conducted as deemed necessary by the project team along the innovation journey. These self-managed evaluations could occur for a variety of reasons such as (Torres-Padilla et al., 2015):

1. To evaluate and select from a variety of choices that may be prepared as a list of ideas and/or presented as 2D prototypes, models, designs and/or roadmaps;

2. To evaluate the desirability, feasibility and/or viability of the business/solution concept based on learning and knowledge gained from experiments (such as thought experiments; 2D, 3D or 4D/live prototypes tests; or full ‘market-ready’ pilots) and make decisions on what would follow, which could include:

   a) Additional research and/or experiments if ‘fit’ is still uncertain (e.g. fundamental unknowns or assumptions remain);

   b) ‘kill’, ‘put on hold’ or ‘pivot’ the project/concept if ‘no fit’ is evidenced; or

   c) move it to the next stage if ‘fit’ has been accomplished, point at which a formal evaluation and approval at project portfolio level might be required
in order to reassess its strategic relevance and commit new resources (e.g. by senior management).

At the project portfolio level, evaluation may be conducted periodically at a convenient time, to assess and prioritise all projects, and focus on the ones with the most potential (i.e. the best portfolio investments) and decide whether a specific project should be moved forward and what resources to allocate to it. At this level, projects are assessed relative to other projects (even to the projects that are ‘on hold’) and on their influence on the overall portfolio. A concept/project should be moved to the next stage when there is evidence of the pursued fit, otherwise, a project should be ‘killed’, ‘put on hold’ or ‘pivoted’. For instance, if new knowledge from experiments undermines a crucial business model assumption, this could result in the decision to ‘kill’/‘put on hold’ the concept/project entirely or it could alternatively result in a ‘pivot’ (i.e. loop back to a previous stage) in order to alter the direction of the project (e.g. in the search of new customers with significant and unfulfilled problems/needs where the solution can fit and create value for them and the organisation).

Evaluation activities are still advantageous once an innovative solution has been formally introduced in the market, and can be used in order to determine whether the value proposition remains valid and if it is being captured for the firm according to the expectations (e.g. revenue growth rate and profitability margins). If the revealed results are not satisfactory, actions should be triggered to conduct one of the following:

1. optimise the operation;

2. improve the value proposition or business model elements by for example pivoting in order to improve the business model for incursion in the mainstream market; or

3. retire the solution from the market, for example when its lifecycle has come to the ‘decline’ phase and the solution itself has proven to no longer be adequately profitable (Torres-Padilla et al., 2015).

The CDA framework aims to evaluate every single concept/project as required, and at the same time consider an entire portfolio of projects to prioritise the projects and focus resources on the ‘best bets’ in order to maximise the potential value for the organisation (Cooper, 1994; Cooper et al., 2002b; Mitchell et al., 2014). Therefore, the aim of the evaluation subsystem is to determine which concepts/projects to
Field studies: case study analysis

devote time, effort and resources to and enable prioritisation activities that lead to investment focus, responding to questions such as the following:

- Risks: what are the possible challenges? What can assist the firm in managing the risks? What is the decision-making criteria?
- Selection: what are the most suitable ideas and concepts? What concepts/projects can provide an appropriate and viable solution to what is occurring or will occur in the environment, and therefore increasing the probability of survival and growth of the firm over time?

To answer questions similar to the above, the evaluation subsystem performs a variety of tasks and utilises the divergent-convergent thinking, where during divergence, the main activities include collecting background information on the project(s) and predicting risks/challenges and/or potential values or benefits. In convergence on the other hand, the main functions are evaluation and prioritisation of concepts/projects in order to communicate, make decisions, and define the next evaluation criteria. Therefore, the mode of thinking during divergence should be mainly oriented towards information collection while permitting the exploration of negatives and positives when appropriate; whereas during convergence expert judgement and ‘gut feel’ is the predominant mode of thinking. Experience and intuition are critical in complex situations (e.g. evaluation of potential innovations), and do not necessarily dictate a lack of rigour (Dissel et al., 2005).

In practice, this subsystem can be executed with the aid of tools and techniques such as the following (Torres-Padilla et al., 2015):

- Scoring methods (e.g. Mitchell, Phaal and Athanassopoulou, 2014) for the aim of evaluating early stage innovation projects.
- Opportunity-Feasibility matrix (e.g. Mitchell et al. 2014; Farrukh et al., 2014) for the aim of prioritising and selecting innovation projects.
- Financial methods (e.g. Cooper et al., 1997) for the aim of evaluating late stage innovation projects.

Tools and techniques

In CDA, procedures - which involve a set of ‘steps’ (or ‘micro-sequences’) configured to achieve outcomes efficiently - are utilised as a means of operationalising functions
through concrete tasks/activities of perception, thinking, decision-making and action. In a collaborative environment, such as workshops, given the constraints in time, people and other resources, flexible steps may be utilised to increase the cognitive, social and emotional benefits. These steps comprise of tasks/activities such as clarifying goals, performing a function, sharing ideas and feelings, reflecting on the experience and outcomes, capturing learning outcomes and information/knowledge gaps, and making decisions on the next steps and responsible people (Torres-Padilla et al., 2015).

In each procedure, tools and techniques are used - where a technique is a structured way in which a function or task/activity is performed - and several complementary techniques may be utilised to instrument a procedure. Examples of these techniques are general purpose techniques such as facilitation techniques used in workshops (e.g. Clustering (Tassoul and Buijs, 2007) for interpreting information, Six Thinking Hats (de Bono, 2000; Myers, 2012) to enable focus and efficiency of collaboration, Storytelling (Hensel, 2010; The Ariel Group, 2011) for engaging people), or more specialised techniques such as Roadmapping (e.g. Phaal, 2010) for aligning markets, strategy, products and technology.

Moreover, a tool supports decision-making (Kerr et al., 2013) by facilitating the practical application of one or several techniques (Shehabuddeen et al., 2000) that aim to connect technological resources with organisation objectives (Keltsch et al., 2011). During workshop-based activities, tools are often in the form of structured templates or wall charts that participants interact with using repositionable notes and pens (Kerr et al., 2012) and are accompanied by simple electronic representations (e.g. presentation slides, word processing documents and spreadsheets that are projected or distributed as handouts). Therefore, a tool enables techniques to become tangible and perform functions guided by a procedure that directs participants towards the subsystem’s desired outcomes (Torres-Padilla et al., 2015).

5.3.2.1 Empathy and framing

As discussed by Brown and Wyatt (2010), many initiatives fail because they are not based on the client’s or customer’s needs and have never been prototyped to solicit feedback. Even when field work is conducted, preconceived notions of what the needs and solutions are, typically create bias. This flawed approach is extremely common in both business and social sectors.

Many authors emphasize that understanding a market in terms of its specific problems/needs is critical in achieving ‘problem-solution fit’. Osterwalder et al. (2014),
for example, states that to understand a market segment, specific customer profiles should be used in order to identify the most important, painful and essential problems/needs and consequently be able to devise ‘great solutions’ (value propositions). Andreesen (2007) also places significant importance on the market (and therefore the associated problems/needs) when he discusses the contribution of the market to the success of a start-up when compared to the product and team quality, and states: “In a great market – a market with lots of real potential customers – the market pulls the product out of the startup. The market needs to be fulfilled and the market will be fulfilled, by the first viable product that comes along”. Maurya (2012) defines the term ‘problem-solution fit’ for the initial stage in his lean startup approach as follows: “...is about determining whether you have a problem worth solving before investing months or years of effort into building a solution. While ideas are cheap, acting on them is quite expensive”. According to Maurya (2012), when the primary questions at this stage have been identified, a minimum viable product that represents a minimum set of features is utilised to allow the right problem(s) to be derived and developed.

At this stage, it is more important to obtain a deep understanding of a prospective customer’s problems rather than the technological details or features (Probert et al., 2013) in order to be able to get beyond the assumptions that block effective solutions, and address the needs of the people who will consume a product or service, and the infrastructure that enables it.

According to this, CDA, as a systematic and collaborative approach for identifying and creatively solving problems using data analytics, includes two fundamental phases of identifying problems and solving problems. These phases are both crucial, however, in practice most people and project teams within organisations are more likely to focus on the latter, which is solving problems. This is mostly due to the fact that we are naturally creative beings, and given any problem, even when largely ill-defined, most of us have a tendency to generate several ideas, that are often not ideas which are both original and that solve the problems with the greatest potential.

The purpose of the first mode of Empathy is to discover new customer insights. One of the challenges for many product development teams specially in data analytics projects, is that they are immersed in the world of products and technologies, which while it is undeniably a critical expertise, it can limit their field of view and perspective because market information typically gets framed in terms of product specifications relevant to existing products. Consequently, well-intended research, even when conducted with product users, is often unintentionally biased toward relatively minor modifications to existing products. Therefore, an effective way to
increase the probability of achieving breakthrough ideas, is to begin with an open exploration of customer needs—especially latent, undiscovered needs that may be difficult to articulate—also referred to as customer insights.

While the purpose of the empathy mode is the development of an expanded understanding of the customer—their thoughts, feelings, experiences, and needs, the purpose of the Frame mode is characterized by a distillation of customer insights and framing of specific insights as well-defined problems to solve. At this point in the process, the team should have an inventory of synthesized information about their customers and their contexts, and the challenge is to identify the needs and insights most worthy of pursuit through the next phase of the process (Brown and Wyatt, 2010).

In this stage, these needs and insights are typically framed as discrete “problem statements” to be used in the next phase as a basis for idea generation, which is the initial activity within the Ideation mode. The team is required to converge on a subset of these problem statements that will subsequently be addressed in the Ideation mode and often multi-voting is used to facilitate this convergence. There are various methods to vote for ideas and/or problem statements, however, the intent is to utilise the evolving wisdom of the group that has collectively benefited from participation in the Empathy mode.

5.3.2.2 Ideation

Truly innovative ideas challenge the status quo, are creatively disruptive and provide an entirely novel solution to a problem many people did not realise they encountered. Linus Pauling, scientist and two-time Nobel Prize winner, argues that “To have a good idea, you must first have lots of ideas.” However, more choices often result in more complexity, especially in terms of budget control and timeline monitoring, and therefore many firms restrict choices in favour of obvious and incremental ideas. However, although this tendency may result in more efficiency in the short run, it typically causes the organization to adopt a conservative and inflexible culture in the long run. Divergent thinking is the enabler and not the obstacle, to innovation.

To enable ideation, interdisciplinary teams often utilise a structured brain-storming process, and by taking one framed question/problem at a time, the group may generate hundreds of ideas. Each idea can be written on a Post-it note and shared with the team, as visual representations of concepts are encouraged, because this commonly assists with others understanding complex ideas and building on them.
Field studies: case study analysis

One rule during the brainstorming process is to defer judgment as it is vital to discourage anyone taking on the often obstructive, non-generative role of devil’s advocate, as Tom Kelley explains in his book *The Ten Faces of Innovation* (Kelley, 2005). Instead, participants are encouraged to generate as many ideas as possible.

5.3.2.3 Prototyping and testing

The goal of the prototyping mode is to better understand the need that a potential market is facing and the ability of an envisioned solution to fulfill that need. Although at this stage, evidence is mainly aimed at market desirability aspects of the concept, preliminary evidence of solution feasibility and business viability can also be obtained.

It is essential to identify, test and validate the assumptions of critical importance as early as possible. Therefore, this stage is primarily about experimentation as an early mechanism to mitigate the risk of considerably wasting resources later by developing a ‘solution’ that is not desired or needed, which might lead to costly and time-consuming activities to solve the situation.

At this stage, demonstrators are utilised to facilitate the evidence of fit and can take various forms to obtain the necessary feedback and responses (e.g. observed behaviours) from major stakeholders. Initially, the purpose of the feedback is not to merely validate but is to facilitate learning. Users/customers often cannot articulate or know what they desire beyond current experience and practice (Ries, 2011; Cooper *et al*., 2002), and therefore, demonstrators with the appropriate degree of richness (i.e. the extent of physical representation, scope, refinement and interactivity (Bradshaw, 2010)) would enable significant insights/evidence to emerge. For instance, low-fidelity demonstrators (e.g. rapid and inexpensive 2D or 3D prototypes) would typically be expected to fulfil the purpose for newly envisioned concepts at this stage, since the goal is not to test specific features, functions and/or appearance of a potential solution but rather facilitate understanding of the real problem/need and its importance to the stakeholder (user/customer). Conversely, higher fidelity demonstrators (e.g. with the appropriate interactive capability) might be appropriate for certain technologies that have already been developed to some level of maturity (possible applications/problems are being explored) (Torres-Padilla *et al*., 2015). Bradshaw (2010) argues that “most authors agree that low-fidelity demonstrators/prototypes are more suitable during the early stages when the emphasis is on understanding requirements and specification development, whereas high-fidelity prototypes are thought to be required for more rigorous testing and uncovering usability problems in the later design stages. The exception appears to be designs
that include a physical manipulation or interaction where some element of feel is involved.”

At this stage, demonstrators that aim to facilitate the evidence of fit, are typically in the form of 2D lean and rapid prototypes supported by workshop-based management tools (e.g. structured wall-charts), which obtain the required knowledge and responses from main stakeholders. A primary set of standardised tools/toolkit such as the popular and extensively tested S-Plan (‘Strategic Landscape’) approach for road-mapping (Phaal et al., 2007), might be utilised as a basis, and other tools might be integrated as needed.

At this point, most of the components explained by the demonstrator are typically extremely raw and an accurate ‘picture’ of details such as market size, may not be possible due to the uncertainty involved, and therefore, qualitative approaches would be more suitable for evaluation at this stage. At this point, market-related factors might be highly uncertain and impossible to assess even in cases where the concept is on the basis of an existing technology, where feasibility can be measured by facts.

As an example, when a concept is based on a potentially disruptive technology, a business model depiction (i.e. a low-fidelity ‘demonstrator’) might contemplate a new and relatively small market, betting on a future entrance into the mainstream market when the technology is further developed (Christensen, 1997), all of which represents a highly uncertain landscape (Torres-Padilla et al., 2015).

5.3.2.4 Data analytics

The purpose of the stages in the data analytics mode is to develop solution’s features and functionalities, as well as the necessary capabilities (e.g. software production and service delivery), in order to create, deliver and valid the user/customer benefits (i.e. the value proposition). The development is conducted in short-time increments in order to enable the market and business assumptions to be constantly tested and alterations or refinements to be made nimbly (i.e. experimentation and development are done concurrently). Product-market fit occurs when there is clear evidence that the value proposition of the solution (built in actual products and/or services) can deliver value to a set of customers that represent a market (Torres-Padilla et al., 2015).

The term ‘product-market fit’ – which is frequently attributed to the high-profile entrepreneur and investor Marc Andreesen- is extensively used, especially in literature associated with lean startup approaches (e.g. Ries, 2011; Maurya, 2012; Blank, 2013). According to Andreesen (20070, ‘product-market fit’ is defined as “being in a
good market with a product that can satisfy that market”. In his ‘startup pyramid’, Ellis (2009) places ‘product-market fit’ as a precondition for effectively scaling a business. The term was further popularized by the Lean Startup movement, which was initiated with Steve Blank’s approach for startups (Blank, 2013), where he describes the value of his methodology to achieve ‘product-market fit’: “[It] proves that you have found a set of customers and a market who react positively to the product: By relieving those customers of some of their money”.

Ellis (2009) on the other hand argues: “Product/market fit has always been a fairly abstract concept making it difficult to know when you have actually achieved it. Yet many entrepreneurs have highlighted the importance of creating a product that resonates with the target market”. As some scholars argue (e.g. Rogers, 2003; Moore, 2014), the probability that the initial set of customers will be originated from the technology enthusiasts and early adopters categories increases as the novelty of the development rises. Technology enthusiasts (also called the innovators) are often the initial customers to appreciate and adopt an embryonic innovation for its own sake, and early adopters (also called the ‘visionaries’) have the insight and temperament to connect the innovation to a strategic opportunity; and therefore, both are willing to adopt an unrefined/incomplete ‘solution’ and support its evolution (Torres-Padilla et al., 2015).

In the CDA framework, the evidence of fit will be facilitated by the use of demonstrators, such as physical 3D or 4D/live (real time) prototypes or minimum viable products that continuously increase in fidelity (e.g. allow to experience the functional and emotional features). For instance, evidence that users/customers are having a pleasant experience may be measured by qualitative feedback but most significantly, by quantitative data such as users coming back and/or paying for a (demonstrator) product/service.

As stated by many authors, as demonstrators/prototypes are developed and tested, so does a product/service and capabilities. According to Miller (1995) “participatory research with stakeholder testing of prototypes is coupled with a new business process that concurrently develops capability and architecture [including product platforms]”. Bradshaw (2010) further adds to this view and explains that “physical prototype development also helps to build new capability by providing a mechanism to simultaneously test organisational capability, market assumptions and technical feasibility, thereby allowing a new form of concurrent learning to develop”.

The CDA framework provides a means to effectively and efficiently convert a concept into a market-ready innovative solution by incorporating iterative development with
short time-boxed increments to demonstrate a concept to stakeholders early, fast and often; allowing activities to overlap as required (Cooper, 2014; Blank, 2013; Ries, 2011). The goal is to gradually develop and learn via deliverables, capabilities and prototypes that become increasingly richer (Liedtka and Ogilvie, 2011) and therefore, create a product/service based on facts about the desirability, feasibility and viability of a solution (IDEO.org, 2015) rather than executing a marketing and sales plan (Blank, 2013). This is in order to rapidly prove the existence of a market that reacts positively to a certain solution’s value (i.e. a set of paying customers or frequent users), before committing and scaling resources prematurely in the wrong route to market (i.e. wasting them). Therefore, the purpose is to efficiently achieve ‘product-market fit’, or in other words, evidence that the product and service offerings are indeed creating customer value and are obtaining traction in the market (Osterwalder et al., 2014; Torres-Padilla et al., 2015).

Consequently, the aim of this subsystem in CDA is to learn about the value of a solution for potential users/customers as well as other stakeholders such as suppliers and employees, by accelerating the transformation activities that lead to value creation for customers, and responding (as best as possible) to questions such as those suggested by Torres-Padilla et al. (2015): what is feasible to be developed? What attributes or features of the solution are valued most by the user/customer (if any)? How positive do they rate their experience with the solution? Are they returning? Are they willing to pay in any revenue scheme? What should be sourced internally and what externally? What are the potential suppliers/partners, capable of delivering, in terms of the required quality and scale? Are they prepared to partner under certain financial (and other) conditions? What is viable that fits within the organisation’s strategy? What should be the marketing and sales strategy, process and resources for a successful execution? In summary, what should be the business model over time that fulfils the stakeholders expectations?

To answer these types of questions, the subsystem, in addition to the routine tasks in the analytics life cycle described in some detail in section 2.3.1., performs a variety of functions, where during divergence, the main functions comprise of designing ‘live’ or market-ready deliverables or prototypes, and full pilots, whereas in convergence, the main functions comprise of planning/defining, developing and measuring the outcomes in order to facilitate learning and decision making. Therefore, the mode of thinking during divergence should be primarily positioned towards experiment design; whereas during convergence the mode of thinking is diverse. For example, combining intuition with accurate information (e.g. facilitated by roadmap and concept/project brief charts) may be valuable in defining/refining a plan for an experiment, that in turn could be used to develop a prototype and test/measure as-
sumptions based on data. Hence, in this way, the intellectual effort, can be efficiently
directed towards the desired type of knowledge outcomes: facts about solution (e.g.
product/service value proposition, technology and capabilities) and market (Torres-
Padilla et al., 2015).

In practice, this subsystem can be implemented with the support of tools and tech-
niques such as the following:

- Agile development methods (e.g. Satpathy, 2013) to manage the development
  of potential innovations in a changing and uncertain environment.

- Make or buy methods (e.g. Probert, 1997; Canez et al., 2000) to develop the
  make-or-buy strategy.

- 4D/Live prototyping (e.g. Liedtka and Ogilvie, 2011; IDEO.org, 2015) to de-
  velop and demonstrate an innovative component or system, such as a product
  or service, in real world conditions with stakeholders (e.g. potential customers
  and partners), and validating and learning about its fundamental assumptions.
6 Evaluation

6.1 Action research

In order to evaluate the CDA framework, I conducted action research using multiple case studies. The assumption of action research is that the researcher cannot be disconnected from social debate and therefore it aims to make the researcher’s biases explicit and record all the successful/unsuccesful interventions (Lyytinen and Klein, 1985). Action research involves not only observing and recording but also actively taking part in endeavours to solve problems on site and essentially intervening in a situation. Therefore, action research is appropriate for the evaluation of the CDA framework, as it allows me to record actions that result in successful or failed outcomes as well as taking an active role in enabling these outcomes by intervening or correcting actions in the project settings.

The action research was conducted in data analytics projects at Commonwealth bank of Australia in collaboration with several different business units of the bank. Time and resource restrictions limited the number of case study projects that CDA framework could be implemented within, to three projects. I did not organise these projects and did not choose the business units that were involved. Two employees of the bank that worked at the Innovation Lab - which is a business unit inside the bank, focused on cultivating innovation - organised and chose the business units that we approached with the CDA idea. These two bank employees where the general leaders of the projects and the key organisers of various aspects of the projects such as recruiting team members, obtaining permission from the bank for various activities and obtaining approval for CDA projects from the appropriate management personnel. I participated in these three projects as the leader of the CDA framework, team member and facilitator. The duration of each project was 12 weeks (3 days/week full-time). CBA has agreed to allow the IP co-created with the bank in my capacity as a leader and facilitator, to be discussed in this dissertation. However, the observation and analysis of the process that is presented in this dissertation is my own work. All co-created material are labelled adequately.
For each project, a team was formed that consisted of 6 core members from various parts and divisions of the bank. Further details about the team are provided in the "Team" subsection.

The aim of these data analytics projects was to create innovative products and services that resulted in considerable business and/or customer value. These projects allowed me to investigate if the probability of discovering high value problems or opportunities that can result in innovative solutions that create significant business and customer value, systematically increases when the CDA framework is used. Since I was unable to obtain financial or long term data on the success of the three completed projects (in general estimates of financial expectations for projects would be unreliable and difficult to obtain, since companies consider them to be confidential), my measure of success of the projects, is limited to the opinion of the managers and experts and how they view the innovativeness of the solutions presented at the end of each project (the ability of these products and services in generating a competitive advantage), the potential of these solutions to succeed (achieve the business unit’s objectives), and also the ability of the CDA framework in systematically increasing the probability of data analytics endeavours resulting in innovative products and services.

In this chapter I will discuss one of the data analytics projects conducted in the Commonwealth Bank, in detail. This Creative Data Analytics project, which was focused on developing innovation, was conducted for CommInsure General Insurance. The two remaining CDA projects - which are not discussed in this chapter - were conducted for two separate business units within the bank (IT Wealth (with the aim of improving the relationship and collaboration between business and IT) and analytics for re-imagining the credit cards space) in a similar fashion to the CommInsure General Insurance project.

CommInsure is part of the Commonwealth Bank Group, with roots in the Australian insurance industry dating back over 138 years. CommInsure provides financial solutions for life insurance and general insurance, as well as investments and is one of Australia’s largest life insurers with $1.97 billion worth of in-force premiums and 15% market share. The next subsections discuss in detail the various stages of the CDA framework conducted for this project.
6.1.1 Pre-requisites

The project was conducted according to the principles of the CDA framework and consequently the first step was to satisfy the necessary pre-requisites of the framework. Each of these pre-requisites are comprehensively reviewed below.

Managerial support:

Managerial support is crucial to the success of data analytics projects and to obtain sponsorship for the General Insurance project, our team delivered a presentation outlining the framework and its principles to the leadership team of General Insurance unit in CommInsure. The leadership team included the Executive General Manager of Wealth, the General Manager of General Insurance and the 12 Executive Managers in the division. Our pitch deck for the presentation included information on the CDA framework, Design Thinking and power of Big Data for innovation. The team secured $120K in funding. To obtain the sponsorship of the leadership team for the project, we were not required to guarantee a certain output or the end result, rather the requirement was convincing the leadership team of the value of conducting data analytics in a creative and user-centric way.

Once the funding was secured and we had the official approval from the leadership team for the project to begin, it was critical to thoroughly understand the values, interests, motivations and emotions of managers responsible for General Insurance and all other stakeholders involved, in order to ensure that managerial support was maintained throughout the project and adequate resources were available. We conducted two meetings with the leadership team to specify the boundaries and the scope of the project. Using the “rope of scope” method, we discussed various topics (such as migrants, and new products to the bank) and specified which topics and areas fell within the scope of the project and which fell outside the scope. This allowed us to determine the areas that the leadership team desired the project to focus on. During these meetings the project brief was created with managers and the business model canvas was the primary tool used to facilitate this. Completing the business model canvas with the leadership team enabled us to understand and empathise with the key stakeholders and managers and understand what was important to the leaders, the assumptions about the business and the business models, etc. Additionally, becoming familiarised with the key assumptions of the business was advantageous because during the CDA process, we were able to test these assumptions in an attempt to determine their validity. Figure below illustrates the business model canvas generated during these meetings.

Team:
Our team comprised of multi-disciplinary bank employees from different units of the bank and with different backgrounds. This was to ensure the team possessed a variety of skill sets and ways of thinking. All team members had an interest in insurance but they were from different divisions of the bank such as Wealth, IT, and Digital. In addition to having a depth of specific skills to allow tangible contributions to the project, all members were ensured to possess the characteristics of being problem solvers, fast learners, interested in innovation and being pro-change. The team members were also chosen based on their level of enthusiasm in being part of a data analytics innovation team, as enthusiasm was crucial to the adequate performance of team members during the 12 weeks project.

Physical environment:

The CDA projects were conducted in a dedicated space called “The Creative Data Analytics Innovation Garage” (CDAIG), which was purposefully designed and prepared to promote collaboration and innovation according to the principles discussed in section 5.3.1.4. Dedicating a special space to the CDA projects was important as
it influenced the perspective of people and placed the shared goals front and centre. The CDAIG provided the team with an energising space to focus, think, act, share, reflect and decide throughout the entire CDA process, and facilitated the development of the required strategic conversations and trust.

Tools were brought to life inside the CDAIG and the use of resources, materials and knowledge (such as pictures, posters, data, ideas, workflows, and prototypes) that were related to the project (and having them always on display within the CDAIG space), enabled the project participants to be continuously stimulated by them. All generated innovation work, would stay in the room and was accumulated over the course of 12 weeks. This not only helped with the creativity of the team but it also allowed the team to track our progress and have a constant visual reminder of the volume of work conducted (which assisted with team members’ motivation). Also, this visual setting was useful as a means of quickly communicating with the managers about the overall progress of the project, the tasks conducted and the findings.

**Team culture:**

Having a culture that promoted experimentation and learning from constant iteration and testing was critical for the success of this CDA project. It was also important for the team culture to be open and collaborative, providing team members with adequate trust and support to enable creativity. In general, the culture of the team was in-line with the design thinking values for cultivating innovation. During the ideation processes for example, team members would not criticise each other’s ideas and viewpoints and there was a “Yes and” mentality in the group, where team members were encouraged to not negatively judge ideas and instead build on each other’s ideas and solutions. Team members would also separate the ideation and evaluation processes from each other.

It was also important for all participants to not have a fear of failure, and be open to experimentation and learning from doing.

### 6.1.2 Challenge identification

From the project brief created with the leadership team, the two most important areas of focus were identified to be acquisition and retention. Consequently, as a team we created four challenge statements around acquisition and retention for the project. Different tools such as the funnel of focus were used to aid us in devising challenge statements that were not overly broad or narrow, were not biased towards a specific solution, and would spark creativity.
The four challenge statements were subsequently presented to the leadership team and they made adjustments and modifications to the statements. It should be noted that having managers not be part of the team and only be present as “gate-keepers” is not consistent with the CDA framework, however we were unable to change this culture in the bank. The final challenge statements are presented in the figure below:

![Challenge Statements Diagram](Image)

These challenge statements acted as a frame for our area of focus for identifying problems and possible solutions.

### 6.1.3 Empathy

The purpose of this mode of the CDA framework was to discover new customer insights and focus on the total customer experience, in contrast to more traditional approaches of feature-based product development and innovation. Our focus was to enable product and service development that was related to customer motivations (what problem is being addressed?) and the benefits the customer extracts through the product/service experience. A deep, empathic, human-centered approach was therefore the critical first step in our data analytics project. As articulated by Tim Brown (2008), knowledge of “human behavior, needs, and preferences” is what helps “capture unexpected insights and product innovation” that will be more desired by consumers.
The aim of this mode was to prevent unintentional bias towards relatively minor modifications, by identifying customer insights and focusing on gaining empathy with customers, that is, developing an understanding of their context, experiences, and behaviours. This was particularly important as one of the challenges for many product development teams is that they are predominantly concerned with products and technologies, which while important, can limit their field of view and perspective, as well as creating a tendency to frame market information in terms of product specifications of existing products.

To enable empathy, we used the previously devised How Might We statements (HMWs), and conducted semi-structured interviews with users to understand their pain-points, problems and needs. The research was qualitative and did not intend to provide comprehensive knowledge on consumption and behaviour segments, and rather to enable empathy and insights from customers/users. In order to determine the specific people we needed to interview, we took advantage of the principle of outliers, which specifies that learning from extremes - identifying extreme behaviours and mapping their patterns, as well as latent needs - enables valuable insights that can be used for average users (but often would not have emerged if the extremes had not been examined). Figure 6.3 illustrates two examples of extreme users in the case of General Insurance.

Figure 6.3: The graph of insurance extreme users (co-created with CBA)
As a team, we brainstormed and ideated about the type of customers and users that could be categorised as extreme users for our project. There was no necessity for the interviewees to be CBA customers. Figure 6.4 indicates the categories of the extreme users we identified for general insurance:

The customer ecosystem map in figure above was created in order to assist us in establishing the type of extreme users we were interested in targeting. The categories of people we identified as outliers and aimed to interview, included:

- Immediate family, parents, or a friend of a person who has, for example, had an insurance policy, a car accident, or have been associated with insurance in some way. This is important in order to allow us to empathise with different perspectives in regards to insurance.

- Uber user: this category may refer to either an Uber driver or Uber passenger. We were aiming to investigate how the drivers and passengers viewed insurance, particularly because at the time of this project many passengers were vary of using Uber as they believed the vehicles were uninsured in the event of an accident.

- Airbnb host: This category was for people that were renting their homes on rental sites such as Airbnb. At the time of this project, Airbnb hosts were
unable to obtain insurance for their property as there was no product that
catered for such users.

- Small business owners were also considered as extreme users as they had chosen
  a different path than the average user and therefore there was a possibility that
  they may view matters differently.

- We were also attempting to target several specific people such as: a miner from
  Western Australia because the WA mining boom had ended, and therefore the
  behaviours of the miners in terms of insurance was interesting to us. Moreover,
  we interviewed a disability pension approver as it was stimulating to compare
  the way they approve/disapprove claims to insurance claims (Parallel Worlds
  method).

As evident from the customer ecosystem map, our aim was to empathise with a wide
range of thought processes and users, and not just with CBA customers.

It was also critical to interview sales and support staff (internal staff of the bank)
that worked in CommInsure or were a supplier of CommInsure. Interviewing this
group was crucial as it allowed us to learn about their behaviours and the way they
operate, which in turn provided us with the necessary context about the behaviour of
the customers. Interviewing the staff also assisted us in preparing more appropriate
questions for the customers, because it facilitated our learning about the internal
processes and the general business logic. At the beginning of the interviews, this
group were treated like other customers/users (were asked the same questions) and
subsequently we proceeded to ask specific staff question. The sales and service
ecosystem (figure 6.5) illustrates the categories that were targeted for this group.

Once all the categories of users that we aimed to target in the interviews, were fi-
nalised, we used an agency to outsource the selection of these people. The agency
identified and selected interviewees based on the detailed selection criteria we pro-
vided them, that not only included the categories in the ecosystem maps, but also
included other information such as age and sex requirements.

The interview questions were created by our team in two, 3 hour long workshops in
which all the members of our team participated. We devised 40 interview questions
from which we eventually selected 20 as the final version of the interview questions.
The aim of the interview questions were to obtain deep understanding of specific
events, feelings, and needs that extreme users had experienced. The interview script
was not always followed exactly and was mostly used as a guide.
The first few questions we asked the interviewees, was always around personal information such as their professional background and family situation. The reason for this was twofold: 1) learning more about the interviewees specific background that would help us with categorising the interviewee in the later stages 2) This would make the interviewees feel more comfortable and would create a rapport with them.

The customer interview questions are given below:

After each interview, as a team, we would conduct a debrief and discuss the key findings from each interview. The interesting or unexpected insights would be written on the board and we would then investigate if any of the other interviewees had similar insights, behaviours or experiences. The aim was to determine whether those behaviours and attitudes were common or if it was a one-off experience. Our questions were specifically chosen not to lead the interviewees towards a specific answer and be biased.

Furthermore, after conducting each interview, the interview audio was transcribed and the project team began synthesizing it (where data synthesis is the process of summarizing and deriving meaning from the data). This did not mean that the empathy work and discovery were complete, as the discovery mode is built on iteration between data collection and data synthesis. Because the collected data...
was qualitative in nature (i.e., pictures, transcripts, audio recordings, etc.), the data synthesis process was considerably different than what is often assumed with market research. Instead of working with numerical data and statistics, our team was...
required to translate these qualitative data into specific customer insights. There are various ways to facilitate this, including: coding transcripts, drafting personas and empathy maps of archetypical customers, and journey maps that describe the customer’s current or ideal experience.

While there were various techniques available for synthesizing data, a vital principle was to continually iterate between data collection and synthesis, and attempt to synthesize insights throughout the project as opposed to waiting for all data to be collected. This required flexibility and persistence, however, it assisted in ensuring that the most appropriate methods were utilised as required, as opposed to rigidly prescribing exactly how the research will be conducted at the beginning of the project.

In our project, one of the ways utilised for synthesising the data, was creating an empathy map for each customer interview (after transcribing the interview). An empathy map is a tool for the synthesis of information through visualization of what an interviewee says, does, thinks and feels. This makes it possible to organize the data of the empathy phase in a way that provides an understanding of situations arising from the context, behavior, concerns and even the aspirations of the user/customer (or other agents subject to examination).

From the empathy maps, we made point of views for each customer. Instead of creating a HMW for each interviewee and ideating based on those HMWs (we interviewed 60 people and therefore there would have been 60 HMWs), we chose to create personas. In the past decade, personas have received considerable attention from both academics and innovative companies in product design (Blomquist and Arvola, 2002; Chapman and Milham, 2006; Faily and Flechais, 2011), due to a positive trend toward creating user-centric products or services. Personas provide designers with a user-centric tool that represents an ideal user (Cooper et al., 2014) that assists designers in maintaining focus on the ideal user as they explore and develop solutions.

Personas are fictional characters that are a representation of ideal or prototypical end users, which are created from a synthesis of observed behaviour among users/customers with extreme profiles (i.e. the people we interviewed) (Cooper et al., 2014). They signify motivations, desires, expectations and needs, combining the important functionalities of a more comprehensive group and represent clusters of users from research and not derived from stereotypical assumptions (Cooper et al., 2014). Personas enable designers to relate to and empathize with users, and encourage them to view product problems from a user’s perspective.
Based on data from the field (interviews and observations), different polarities of user characteristics were identified. Typically, these polarities may include a range of characteristics from demographic features, such as gender, age group and social class, to behaviour profiles (for instance, if the individual is independent with regards to finances, or if they depend on family members to manage their finances and insurance), however, we did not group interviewees based on demographics or age and instead the categorisation was based on their attitude towards insurance. After identifying all polarities, the characters were constructed by combining these features and using the profiles identified in the field as reference points. Therefore, a group of 5 personas was created with considerably different characteristics representing extreme profiles of users of General Insurance. A name was assigned to each persona and stories and needs were generated to assist in the “personification of each archetype.”

The personas created are shown in figures 6.7 to 6.11. Every person we interviewed could have been placed in one of these persona categories indicating that they had the same thought process/attitude towards insurance.
Another primary method used in our project to understand the total customer experience was journey mapping (also called experience mapping), which assisted in
understanding, synthesizing, and forming insights about the total customer experience. Experience mapping is part of many design thinking toolboxes (e.g., Fraser,
Evaluation

2012; Kumar, 2013) and is directly linked to other methods in the design process such as personas, ideation, and stakeholder value exchange. Based on the insights from the interviews, we created the customer engagement journey shown in figure 6.12, which depicts the life cycle and experiences of CBA potential customers with various insurance products. The goal was to create an experience-based facilitator for product design and innovation.

The experience journey revealed that most renters were not even aware that the bank provided insurance products suitable for them. Therefore, as depicted in figure 6.12 renters were in the pre-awareness phase and they did not purchase from the bank. Similarly, car owners were not aware of the fact that the bank offered car insurance and they did not associate insurance with a bank. Once they became aware that the bank offered car insurance however, the pricing and risk assessments of the bank acted as a critical block, preventing customers to proceed with the purchasing journey. If the customers were able to get through this hurdle however, they fluctuated between the evaluation and committing stages.

For home insurance products on the other hand, if the customers had their home loan with the bank, they would commit to a home insurance without passing through any other parts of the engagement journey. This was due to the fact that they had received the home insurance as part of their loan package and consequently, they
Evaluation

did not need to be aware of the offering or to evaluate it, they simply obtained the product due to their loan product.

6.1.4 Framing

In contrast to the Empathy mode, which is concerned with the development of an expanded understanding of the customer — their thoughts, feelings, experiences, and needs - the Framing (Define) mode is characterized by a distillation of the collected customer data and framing of specific insights as well-defined problems to solve. By this point in the process, the team had an inventory of synthesized information about the customers and their contexts, and the challenge was to identify the needs and insights most worthy of pursuit through the next phase of the process. From each interview we had gained valuable insights and some of these insights are shown in figures 6.13 to 6.17:

Figure 6.13: Key insights from empathy phase - Word of Mouth (co-created with CBA)
Evaluation

Positive referrals can turn into

**Word of Mouth**

What our users said...

- Natasha: “If I know the brand and it is reliable, I will pick them over price”
- Fiona: “I told my friends of my Youi experience”
- Hannah: “Once I am loyal, I’m a super advocate”
- Steve: “When I have a bad or good experience I tell all my friends about it”
- Hannah: “I am big on word of mouth. You always hear stories”

Figure 6.14: Key insights from empathy phase - Word of Mouth quotes (co-created with CBA)

People don’t think of Insurance when they go into a bank branch.

When they think of insurance they think of online.

What our users said...

- Colin: “Banks do not do insurance”
- Mandy: “I don’t associate insurance with a bank”
- Fiona: “I don’t think insurance when I go into a bank branch”
- Rebekah: “Branch is time. Time is precious. Waiting in a line is a waste of my time”

Figure 6.15: Key insights from empathy phase - Insurance not associated with banks (co-created with CBA)
Banks do not sell Car Insurance.

Fiona
“Cars are a very specialised service. Banks are for home and contents only”

What our users said...

Parramatta
“Car insurance is such a hard sell”

Vitally
“I told my friends about my [CBA] car insurance. They thought it was unusual”

Lorraine
“Banks don’t do car insurance, insurance companies do”

Figure 6.16: Key insights from empathy phase - Car insurance not associated with banks (co-created with CBA)
Typically, these needs and insights are framed as discrete “problem statements” to use in the next phase as a basis for idea generation in the ideation mode. The team was required to converge on a subset of these problem statements to address in the next mode (ideation), and multi-voting was frequently used to facilitate this purpose. While there were various ways to vote in order to converge on a few problem statements, the intent was to take advantage of the evolving wisdom of the group that had collectively benefited from participation in the Empathy mode. Using the personas and insights from the customer/user interviews we created 6 opportunity spaces, which are depicted in figure 6.19:

Below, I will discuss each of these opportunity spaces in more detail:

Awareness: As it was revealed through the empathy stage, customers were unaware of CBA’s insurance products - especially the car insurance products - and therefore there was a tremendous opportunity for the bank to increase customer awareness in regards to the available products. Furthermore, the majority of the migrant population in Australia banks with CBA, however they rarely have insurance products either because they are unaware CBA offers it or because they are unfamiliar with insurance products in general (they might have not even had such products in their country). Another surprising factor was that the CBA staff did not know CBA
offered insurance. Therefore, raising awareness about the insurance products was a significant opportunity space.
Word of mouth: In the initial stage of the project when we met with the leadership team to define the challenge, we asked them to walk us through the customer journey map. From the perspective of the business, the customer journey started with research about the insurance product. However, through empathising with the customers we deduced that the first step is indeed not research and is instead a whole missing step that we call “word of mouth”. Customers would approach trusted members of their family and network for advice regarding insurance and would typically settle for the same products as them (specially in the case of children/parents). However, the bank was not considering this important first step in the customer journey effectively.

Transparency: This opportunity space was concerned with the customers’ lack of trust in insurance providers. Many customers believed that the insurance providers were not aiming to help customers and were rather only interested in charging the clients and avoiding claims payments by utilising their extensive and hard to understand terms of contract. This opportunity space was concerned with being transparent with customers and gaining trust back.

Advocacy: this opportunity space was about determining ways of encouraging customers to advocate for the bank’s insurance products. Examples were parents who had current insurance policies with CBA, acting as advocates and influencing their children to also obtain insurance products from CBA that they deemed appropriate.

Connection: Most customers believed that because they have been paying a premium and have been loyal to the bank, they deserve to be rewarded. Lack of reward resulted in customer dis-satisfaction and reduced loyalty, and therefore this opportunity space was concerned with connecting the customer and the bank.

6.1.5 Ideation

For each opportunity space, we created multiple HMW statements that were used during the ideation sessions. In total, we held 18 ideation sessions (one for each HMW statement) that resulted in 1639 ideas. In each ideation session, in addition to the core team, 12 completely new CBA employees were present. The employees were chosen to be from a diverse range of backgrounds and business units.

Before each ideation session, we sent an “ideation pack” to all new participants, informing them of the general purpose of the ideation session as well as providing them with the necessary background information for the project (e.g. some of the insights and opportunity spaces). This enabled the participants to have a prior understanding of the purpose of the session and begin to think about the project
unconsciously or intentionally. This would also prevent us from having to spend a large amount of time going through the background information in depth during the ideation sessions. At the start of each session, a high-level overview of the project was provided to participants in order to refresh their memory.

Use of different tools also assisted with the ideation process because without tools people have a tendency of choosing or creating safe options. Our aim was to create an environment that made participants feel comfortable with ideation, and many techniques and tools such as parallel worlds, free brain storming, and break the worlds, were utilised.

Figure 6.20: How Might We questions for general insurance (co-created with CBA)
Subsequent to grouping and refining ideas, the team used various forms of multi-voting to converge on the most promising ideas, by considering a broad set of criteria.
that included (a) desirability (from the customer’s perspective), (b) feasibility (the ability to deliver the product), and (c) viability (the ability for sustained business benefit, either financial or strategic). It was critical, however, to remain focused on the identified customer insights and to avoid filtering ideas excessively based on other criteria, since the idea was still emerging at this point and could be improved upon during the next activity, prototyping.

After multi-voting, ideas with the most votes were grouped in the categories shown in figure 6.23 below.

![Concepts were grouped into 4 categories](image)

**Figure 6.23: Grouping the concepts (co-created with CBA)**

### 6.1.6 Prototyping

The aim of the prototyping mode was to develop a concept or set of concepts that can be shared with the target market for feedback and that, through iteration, can be improved upon. Although customers are able to respond to an idea on its own, the best feedback will come from engaging with a rough prototype of a concept because a good prototype can provide an experience to respond to and another opportunity for the team to observe actual behaviors.

In this stage we used simple prototypes – often referred to as “low-resolution prototypes” - that provided basic experiences of products or features of a product (d.school, 2014). These early-stage prototypes were three-dimensional objects, a
sequence of screen shots of a “software app” concept, or a mocked-up service counter with actors as agents.

The purpose of this mode was to obtain feedback on concept prototypes, and the ideas and assumptions embedded within them. In the CDA framework, much of the feedback was used to iterate and improve upon the concepts, especially in the first iteration of the first four modes. In other words, this was not the “final step” before the development and testing and iterations were continuously occurring. Prototyping was predominantly used as another activity for exploring an idea to accelerate and improve idea generation by considering different manifestations of the concept. Consequently, it was important to appreciate that the purpose of the feedback was initially as a mechanism to learn more rather than merely to validate.

Figures 6.24 to 6.35 illustrate several of the prototypes developed.

---

**Figure 6.24: Sample prototype - Offering retension incentives (co-created with CBA)**

**CATEGORY: Product Re-design**

**Leaving customers – offering retention incentives**

**Background:**
- Customers who’ve never made a claim are typically swayed by price. Once customers have claims experience, then they are typically swayed by the level and quality of service. Their claims experience also heavily guide their retention decisions.
- Customers told us they will research 2-3 weeks prior to renewal. As the renewal dates approaches, customers will be less likely to research.

**How It works:**
- A customer cancels their policy and one month later, we send them an offer (via mail or CommBank app), with a discount to a partner store, inviting them to reinstate their policy to receive the voucher.
- The second part to the prototype involves sending a voucher as part of renewal.

**Prototype type**
- Physical wireframes shown to the customer

**Real prototype?**
- Yes / No

**User needs this addressed:**
- Customers want to be recognised for their loyalty.
-various nupness
- Would this offer encourage customers to return back to Commun jews?
Figure 6.25: Sample prototype - Setting up a new policy (co-created with CBA)

Figure 6.26: Sample prototype - Moving to new address for under 30s (co-created with CBA)
Evaluation

Figure 6.27: Sample prototype - Insure your child’s toy (co-created with CBA)

Figure 6.28: Sample prototype - Customisable renters product (co-created with CBA)
Evaluation

Figure 6.29: Sample prototype - Content insurance (co-created with CBA)

Figure 6.30: Sample prototype - Renewal with competitor (co-created with CBA)
Figure 6.31: Sample prototype - Renovation poke (co-created with CBA)

Figure 6.32: Sample prototype - Insure the device for free (co-created with CBA)
Evaluation

Figure 6.33: Sample prototype - No excess (co-created with CBA)

Figure 6.34: Sample prototype - Commit to Fix (co-created with CBA)
6.1.7 Testing

The purpose of this mode was to obtain feedback on the concept prototypes, and the ideas and assumptions embedded within them. Most of the feedback obtained was utilised to iterate and improve upon the concepts, and facilitate learning as opposed to merely validate.

Two types of activities were conducted in this mode. The first activity was to share prototypes with potential customers to gain feedback. To enable obtaining the most valuable feedback, the prototypes were utilised to assist in simulating an experience for the user rather than serve as a prop for presentation. After the team collected sufficient feedback, we proceeded with a process of synthesizing the feedback that was an activity similar in spirit to the data synthesis completed during the Empathy and Framing modes, with the obvious difference being that in this stage users had a tangible solution concept to respond to. The objective however, was similar and it was to gain further insights and converge on the most promising solution or elements of a solution.

For testing of the prototypes we attempted to have most of the users that we interviewed during our empathy work back. Having access to the initial users for testing was desirable as some of the prototypes were created specifically from their feedback
and it was advantageous to regain empathy with them. This was to first investigate if we had understood them correctly previously and whether the proposed solutions met their needs. We also had entirely new users participate in the testing stage as it was important to have fresh perspectives on the ideas as well.

Based on the feedback obtained from testing, the prototype concepts would proceed to one of the paths shown in Figure 6.36: BAU (the idea does not contain much excitement but the customer believes it should already be in place), Persevere (the idea to proceed with minimal (if any) alterations), Pivot (incorporate customer feedback and redesign), Perish (the concept is not suitable as it does not meet customers’ needs).

Figures 6.37 to 6.48 represent the results of the testing of several of the concepts and prototypes.

![Figure 6.36: Paths that prototypes could take based on testing results (co-created with CBA)](image-url)
Figure 6.37: Tested prototype - Setting up a new policy (co-created with CBA)

Figure 6.38: Tested prototype - Moving to new address for under 30s (co-created with CBA)
Evaluation

Figure 6.39: Tested prototype - Leaving customers (co-created with CBA)

Figure 6.40: Tested prototype - Insure your child’s toy (co-created with CBA)
Figure 6.41: Tested prototype - Customisable renters product (co-created with CBA)

Figure 6.42: Tested prototype - Content insurance (co-created with CBA)
Figure 6.43: Tested prototype - Renewal with a competitor (co-created with CBA)

Figure 6.44: Tested prototype - Renovation poke (co-created with CBA)
Figure 6.45: Tested prototype - Insuring the device for free (co-created with CBA)

Figure 6.46: Tested prototype - No excess (co-created with CBA)
Figure 6.47: Tested prototype - Commit to fix (co-created with CBA)

Figure 6.48: Tested prototype - Empty investment properties (co-created with CBA)
Figure 6.49 represents the main achievements of the previous stages.

6.1.8 Data analytics

Lavalle et al. (2011) state in their study that organisations should implement analytics by first defining the insights and questions required to meet the critical business objectives and then identifying the sources of data needed for answers. By this stage in our project, concepts and prototypes that offered the most significant potential for value creation were identified. Defining the desired insights first, was beneficial because we could target specific subject areas and use readily available data in the initial analytics models. This in turn facilitated the discovery of the gaps in the data infrastructure and business processes by exploring the insights delivered through these initial models. Consequently, time that would have been spent cleaning all data was redirected toward targeted data needs and specific process improvements that the insights identified, enabling iterations of value.

Another advantage of defining the business problems first was that if we had made data our overriding priority there was a high probability of losing momentum before the first insight was delivered, because a data-first approach can typically be perceived as taking too long before generating a financial return (Lavelle et al., 2011). Consequently, by narrowing the scope of these tasks to the specific subject areas
Evaluation

needed to answer the main questions and concepts, value was realized more quickly; while the insights were still relevant. Furthermore, not starting with the data or process change allowed us to not encounter unintended consequences — such as data that is not extensible or processes that are ultimately eliminated — that require rework and additional resources to solve.

In this stage, we first proceeded to identify the data sources needed for the development of the prototypes and concepts that were selected in the previous stage. In order to obtain internal data, I investigated the data architecture of the bank as well as limitations/pre-requisites for being able to obtain specific types of data.

Figure 6.50 shows a typical data architecture at Commonwealth bank of Australia. This data architecture presents several challenges for conducting data analytics for innovation.

Figure 6.50: A typical data architecture at Commonwealth Bank (co-created with CBA)
First data analytics for innovation needs workspaces that are purpose-built for experimenting with data with flexible and agile data architectures. However, the CBA data warehouses provide excellent support for traditional reporting and simple data analysis activities but have a difficult time supporting more robust analyses.

In this architecture, for data sources to be loaded into the data warehouse, data needs to be well understood, structured, and normalised with the appropriate data type definitions. Although this kind of centralization enables security, backup, and failover of highly critical data; it also means that data typically must go through significant processing and checkpoints before it can enter this sort of controlled environment, which does not lend itself to data exploration and iterative analytics.

As a result of this level of control on the Enterprise Data Warehouse (EDW), additional local systems have emerged in the form of departmental warehouses and local data marts that business units create to accommodate their need for flexible and in-depth analysis. Accessing data held by different business units is also difficult, and even if the business unit has access to the data, there are limitations on what the data can be used for. These limitations mostly stem from the privacy laws concerning personal information.

The definition of ‘personal information’ from March 2014 extends to information or an opinion about an individual who is reasonably identifiable. Information is considered personal information if the identification or re-identification is practicable from the information itself or in combination with or reference to other information. This is true even if the information is in public domain with the individual’s consent. Deloitte (2013) argues that there are three major subjects that cause the majority of concerns when engaging with analyses on personal data:

- Australian Privacy Principles (APP) 6 outlines the circumstances in which an organisation may use or disclose the personal information that it holds about an individual: If an organisation collects personal information about an individual for a particular purpose (the primary purpose), it must not use or disclose the information for another purpose (the secondary purpose) unless the individual consents to the use or disclosure, or another exception applies (Leonard, 2014).

This means that personal data may be processed for specified and explicitly stated purposes and that these purposes have to be communicated to the consumers. At the time the transaction data was first collected, processing it for value creation may not have been an intended named goal.
• There also must be a formal ground for processing, the reason why using the data for the intended goals is allowed at all. Usually this can be found in the consent of the customers or if the bank that holds the data has a legitimate interest to process it. APP 3 outlines that an organisation must not collect personal information (other than sensitive information) unless the information is reasonably necessary for one or more of the organisation’s functions or activities (Leonard, 2014).

In regards to this APP, Deloitte (2013) suggests that the legal ground be given proper attention and selected carefully, to avoid pitfalls in the future.

• It is important to realise that human decision making should always be part of the process. It is currently forbidden to take automated decisions based on profiling that has a direct influence on a customer – and according to Deloitte (2014) it is unlikely that this ban will soon be lifted.

The above three subjects cause the majority of concerns when working with personal data within the CBA, and these privacy laws set the boundaries for what could be done using analytics and also specified what measures needed to be taken in order to comply with the regulations. However the privacy regulations in regards to personal information are only applicable if the information can result in the identification of an individual. Therefore, if during our project the personal information was transformed in a way that the risk of an individual being identified -by the information itself or in combination with other information- was reduced to impractical, the privacy regulation no longer applied. This transformation could be done through aggregation or anonymisation of the personal information. However, a problem with aggregation and anonymisation of data is that it is time consuming, expensive and not always effective in conserving the privacy of customers. If there is any way in which an individual could be matched and tied to non-identifying transaction data, as Leonard (2014) argues, partitioning of databases would be ineffective in allowing non-identifying transactional data to be used without complying with the rules that relate to use of personal information.

Due to these challenges of performing analytics with personal data and the time consuming and complex nature of data anonymisation, as a team we first needed to decide what data sources we were required to obtain i.e. which data sources had the potential of providing us with the required insights. To facilitate this, we began a process that we called “data storming”. For data storming we simulated data that was difficult to access (due to regulations and/or bank’s privacy concerns) such as transactional data, and combined the simulated data with publicly available
data that we deemed appropriate such as social media and weather data. For the
generation of simulated transactional data that had realistic characteristics, I in
collaboration with IBM Research Australia, created a data generation system that
was used as part of the initial data analytics process to help in the “data storming”
phase and determining the specific data that was required in the project. I was the
lead researcher working on this data generation system and I have permission from
IBM to discuss the co-created IP in this dissertation, in my capacity as the lead
researcher and the leader of the project. The data generator is discussed in detail
in Appendix 4.

After the identification of the data sources, the team started conducting analytics
by obtaining and cleaning the required data and hence creating analytics models.
4D live prototypes were created for each concept and development was conducted
in short-time increments in order to enable the market and business assumptions
to be constantly tested and alterations or refinements to be made as required and
experimentation and development to occur concurrently. Due to the bank’s privacy
concerns, I personally was unable to actively work with the real data in the bank
to conduct data analytics and therefore I was not an active member of the project
during this phase. However, the team did implement the projects and concepts.

### 6.1.9 Interviews with managers

In order to validate the CDA framework’s effectiveness in systematically increasing
the probability of success of developing innovative products and services, I conducted
face-to-face semi-structured open-ended interviews with three executive managers
that were involved in each of the CDA projects conducted within CBA (9 managers
in total). These managers had consistent involvement in each of the CDA projects
from the beginning of each project, and they were also the points of managerial
contact during each project. Therefore, due to their high level of involvement, they
were adequately aware of the results obtained from the projects and how the CDA
framework was used.

The interview questions had the intent of obtaining critical qualitative feedback in
regards to the effectiveness of the CDA framework for developing innovation in the
organisation (see Appendix 3). During the interviews, the Managers were encour-
gaged to compare the results from the completed CDA projects to their previous data
analytics efforts and rate the innovativeness and the potential for success of the out-
comes of the projects. Measure for effectiveness of the solutions was the potential
ability of the solutions to result in an innovative service or product that could lead
to significant business or customer value. The aim was to enable critical qualitative
feedback and comments that could guide enhancements of the framework, rather than rating the framework on a scale of perceived value.

In general, the managers believed that the CDA framework was useful for their data analytics endeavours because it provided a user-centric systematic framework for conducting analytics. They also had positive opinions in regards to the probability of success of products and services that were generated during the three CDA projects. The contribution of the framework was particularly evident when comparing the results of the three projects to previous projects conducted.

In general the problems faced by the bank prior to the implementation of the CDA framework was similar to the problems voiced by the other financial institutions discussed in chapter 2. They indicated that the most significant problem which they faced in their analytics efforts was the difficulty of identifying what problems/opportunities to focus their analytics endeavours on, in order to obtain business and customer value. The managers believed that applying the CDA principles had the potential of creating the ability of deriving innovation from data analytics, by focusing on problems and users and taking on an experimental approach.
7 Conclusion

The purpose of this final chapter is to summarise the research process, discuss the findings and present the conclusions of the research study. Discovering a means of systematically increasing the probability of success of developing innovation through data analytics was the subject of this research effort. In the analysis, the study was based extensively on the field interviews to derive and ground the findings. The aim was not to study every large enterprise that was a leader in the use of analytics for innovation, but rather to look in depth at a sample of such organisations and determine the factors that influence their success in developing innovation through analytics.

The chapter is structured as follows: it begins by summarising the research process, followed by the discussion section that integrates the specific concepts and findings of the field study with insights from the existing theories. Next, the conclusion of the research problem and the contribution of the research are discussed. I then present an evaluation of interpretive research, and discuss the implications of the research findings, limitations of the study and further research.

7.1 Summary of research process

7.1.1 Summary of thesis chapters

In this section, I will present a summary of the main concepts of each chapter of this thesis, which had the main concern of investigating the use of data analytics for innovation and the factors that influence its successful adoption for innovation and creating innovative products and services.

Chapter 1 presented an overview of the ideas and issues that formed the basis of this thesis and laid its foundation. It presented the research problem of the thesis which was the use of data analytics for innovation in large organisations. The study addresses the following research questions: (1) How is data analytics used for innovation in innovative companies (2) What are the factors that influence the
success of use of data analytics for innovation, and (3) How a framework can be created to increase the probability of success of achieving innovation through data analytics. To provide answers to these questions, the study adopted a multiple case study, in order to obtain deep insights and understanding of the phenomenon of the use of data analytics for innovation and its influencing success factors.

The chapter provided justification for the reasons that large organisations need to innovate in order to stay competitive and why the use of data analytics may be the solution for financial organisations. Whilst innovation has always been crucial for the survival of firms, new technologies and digitization are eroding traditional revenues and are creating significant disruptions in the forms of big-bang disruptions and disruptive innovations, making innovation more critical for large firms than ever before. For financial institutions, with no physical products, data – the source of information – is the most valuable asset they possess and one that can be the enabler of unprecedented competitive advantage and innovation. Data contains valuable insights into customers’ preferences and needs, allowing it to be used as a basis for the creation of products and services that remedy emerging problems or opportunities in a timely and competitive manner. It is therefore critical to gain insights into the issues that affect the adoption and successful use of data analytics for innovation in large organisations in general and in financial institutions in particular. Furthermore, an understanding of how innovative organisations use data analytics and the factors that influence its successful adoption and use for innovation, may provide an avenue for creating a general framework to assist other organisations in increasing the probability of developing innovation using data analytics.

Chapter 2 discussed big data and data analytics which form the background of this research. The definitions, use, advantages and potential of big data and analytics for value and competitive advantage creation were discussed. The chapter showed, through primary data collected using qualitative research and secondary data from literature, that while most organisations believe in the power of data for generating innovation and business/customer value, many are struggling to unlock this power and are unable to determine how to conduct their data analytics endeavours in order to create innovative products or services. The chapter discussed current challenges facing large financial organisations in obtaining novel value from data analytics, and presented an overview of the current state of their data analytics capabilities and practices.

Chapter 3 reviewed existing conceptual models, theories and empirical studies relevant to this study. The discussion in this chapter drew from theoretical perspectives of upper echelon theory, Process theory, resource-based view and dynamic capabil-
Conclusion

ities theory, in addition to an extensive body of research in the field of innovation adoption and implementation in large organisations (innovation success factors in general) and innovation frameworks. This review of existing literature and theories allowed me to intensify my theoretical sensitivity, by providing me with insights and the ability to give meaning to data, as well as the capacity to understand and separate critical information from a mass volume of data. It was this theoretical sensitivity that allowed me to develop a theory that is grounded, conceptually dense and well integrated. Furthermore, the existing innovation literature review acted as a point of comparison of the findings of this study, allowing to determine which findings were in-line or in contrast with pre-conducted studies.

Chapter 4 evaluated research methods and described the selected methodology for data collection. Furthermore, philosophical assumptions of positivism and interpretivism were reviewed and it was revealed that an interpretivist view is appropriate for this study, as I aim to deeply investigate a phenomenon that was previously not studied extensively in literature. A framework for guidance on selecting research approaches in IS was presented, and the application of the specific research method selected for this study - which was based on interpretive paradigm and used qualitative case study and action research – was discussed and the justification for selecting this methodology was provided. Research process also obtained considerable attention in this chapter, in terms of both the philosophical foundations of the research method and the practicalities of conducting the study. I provided details of how the data was gathered and how the research was conducted, how the research developed over time, and how the data was analysed. Furthermore, the limitations of the selected research approaches were discussed and the description of grounded theory procedures for analysing data was provided.

Chapter 5 presented the analysis and findings of the conducted multiple case studies. The chapter analysed the large innovative organisations that participated in this study and identified the factors that influenced and inhibited the success of developing innovation using data analytics. The accounts of the use of data analytics for innovation and creation of innovative products and services in the participating organisations were analysed and described using grounded theory technique. The theoretical model is proposed as an initial formulation of the fundamental factors that influence the probability of success of large organisations in developing innovation using data analytics. The case study results have indicated the factors that influence the success of creating innovative products and services using data analytics, as cultural, structural, managerial, experimental, and people related. Based on these identified factors, I also created the Creative Data Analytics (CDA) framework which aims to act as a holistic and end to end solution for strategic data analytics in-
novation management, by increasing the probability of success of creating innovative products and services using data analytics.

Chapter 6 presented the evaluation method for evaluating the effectiveness and efficiency of the CDA framework in enabling innovation through data analytics. Three projects were conducted in collaboration and within the Commonwealth bank of Australia (CBA) and one of these projects is discussed in detail in this chapter. The evaluation method was in the form of action research, where I actively participated in real projects where the CDA was implemented in, in order to not only observe and record but also to take part in the projects to solve problems, intervene in situations, and record all the successful/uns-successful interventions. After conducting the three projects, the CDA framework was evaluated qualitatively, using open-ended semi-structured interviews with relevant managers within the bank and experts in the fields of innovation and data analytics.

7.1.2 Review of research questions

This section reconsiders the research question first presented in Chapter 1 and reflects on the extent to which it has been answered or altered by the study. This research was designed to provide an empirical examination of the use of data analytics for innovation, and to enable the creation of a framework that increases the probability of successful development of innovation through the use of data analytics. In particular, the study investigated the current use of data analytics in the world’s most innovative organisations and developed an enriched model of the factors that influence the success of innovation creation through data analytics based on the findings from theory-driven case studies.

*How to create a holistic end-to-end framework, which increases the probability of creating innovative products/services using data analytics, that result in significant business or customer value?*

This primary research question consisted of the sub-question of:

**What are the factors that influence the success of developing and cultivating innovation using data analytics?**

This question is the main thrust of the argument of the case study phase of this research where the phenomenon of innovation using data analytics is explained and conceptualised, in order provide a comprehensive understanding of the factors that influences the success of data analytics in enabling innovation. Chapter 5 presents the case study analysis, the results and the conceptual model developed from the
Conclusion

The chapter analysed the cases of the innovative organisations that participated in this study and identified the factors that influenced and inhibited the success of enabling innovation through data analytics. The data from the case study was analysed using the grounded theory technique which facilitated the development of our theory which is grounded in the data and is generated from the data in methodical and purposeful ways as a set of common themes that emerged from the analysis.

The success factors from the innovative organisations’ experiences with innovating using data analytics are presented in table 5.3. The tables depicts the core categories and concepts that emerged as critical from the data analysis, as well as their relationships. This theoretical model is proposed as an initial formulation of the fundamental factors that influence the success of developing innovation using data analytics in large organisations. The case results focused on organisational factors which are arguably within the control of the firm, and included managerial, cultural and structural factors as well as people aspects.

The derived success factors facilitated the creation of the Creative Data Analytics (CDA) framework which is the solution to our primary research question mentioned above. The main objective of the CDA framework is to achieve a clear and simple mapping of the identified success factors to stages of the creative data analytics process, in a way that each of these elements can be implemented in any financial organisation (or other organisations) to increase the probability of their analytics efforts resulting in innovation and innovative products/services.

7.1.3 Review of research methodology

Given that this research has a descriptive and exploratory focus, a qualitative strategy which utilised case study methodology was adopted. The study focused on interpreting the context in which innovative organisations currently use data analytics for innovation and the factors that influence the success of such endeavours. The search of literature revealed (to the best of my knowledge) no empirical studies conducted in this area have used a qualitative case study to examine the use of data analytics for innovation in large organisations. The most common method that has been used in studies to examine the use of data analytics for value creation in businesses, has been surveys which has resulted in these studies failing to provide full and comprehensive insights into the use of data analytics and the factors that influence its success for value creation. Previous studies that have used survey method to examine the use of data analytics in large organisations included Lavelle et al. (2011), Olhorst (2012) and McAfee and Brynjolfsson (2012). I argue that
use of survey method alone presents a restricted view of the use of data analytics for innovation and the factors that influences its success. The survey method pays limited attention to the deeper understanding of the factors that influence the success of data analytics endeavours in creating innovative products and services, and therefore no matter how rigorously conducted, it is unable to provide a full insight into this phenomenon. While the survey method is useful due to its ability to handle a large sample size and a large number of variables simultaneously, it is unable to provide a deep view of the phenomenon.

This research is unique in that, to the best of my knowledge, it is the first study to use the case study methodology to examine data analytics enabled innovation in large organisations. I believe that in order to obtain a holistic view of the complex and dynamic phenomenon of the use of data analytics for innovation and its influencing factors, it was necessary to gain a different perspective and utilise the interpretive approach as the paradigm of inquiry. The interpretive case study approach provided the opportunity to conduct rigorous research, which is of direct practical relevance and enabled the discovery of various situations which lead to more knowledge with high potential for relevance for practitioners.

The case study method was used to gain a deeper understanding of the factors that influence the successful use of data analytics for innovation by obtaining qualitative information. It was considered particularly useful in understanding the factors that influence large organisations to effectively use data analytics for innovation, because this phenomenon is developing rapidly and limited information is available about the effective use of the current amounts of data available (big data) for innovation and value creation. The case studies provided a much richer picture of the phenomenon than can be obtained from using surveys (Yin, 1994), which are restricted to gathering information related to the questions on the survey instruments. The use of case study provided a means of collecting a large quantity of rich qualitative data, which was useful in the identification of issues for further investigation (Benbasat et al., 1987).

7.2 Conclusion of the research problem

7.2.1 Case study

The case study findings have provided new insights into the factors that influence the success of data analytics endeavours in developing innovation through innovative products and services. The case data showed that many of the factors that influence the development of innovation through data analytics, are similar to factors
identified for cultivating innovation in large organisations in general. The categories constituting the emerged factors were used to create a framework to increase the probability of achieving success in creating value in data analytics projects.

The results revealed that organisational factors of management support, organisational resources, organisational culture and people aspects, play a dramatic role in the success of data analytics projects generating significant value for business and customers. These results indicated that in addition to technical requirements, success of data analytics endeavours in developing innovation in large organisations is highly dependent on the organisational characteristics and culture in place. The case results showed that innovative organisations that are leaders in the use of data for innovation and value creation, place a significant emphasis on problems to be solved and adopt an experimental and iterative approach to identifying and solving the most valuable problems or opportunities using analytics. Customer and user centricity are also given considerable emphasis in the data analytics endeavours of the large innovative organisations in our case studies.

7.3 Theoretical evaluation of interpretive research

In this section I provide a theoretical evaluation of the interpretive research conducted. It is now extensively accepted to follow the conventions for evaluating information systems case studies conducted according to the natural science model of social science. A set of methodological principles for case studies – that are consistent with the conventions of positivism – were formulated by Benbasat et al. (1987), Lee (1989) and Yin (1994), which resulted in case study research becoming an accepted and valid strategy within the IS research community. However, despite the fact that their criteria are beneficial for the evaluation of case study research that is conducted according to the natural science model of social sciences, their proposed positivist criteria are unsuitable for interpretive research.

It is relatively new to use the interpretive approach in information systems field, and the approach has emerged as a valid and critical component in information systems research where most mainstream IS journals now welcome interpretive research and significant groups of authors are working within the interpretive tradition (Walsham, 1995). A fundamental purpose of interpretive research is to seek meaning in context, which can assist researchers in understanding human thought and action in social and organisational context. This has the potential of generating deep insights into information systems phenomena including the use and the management of information systems. Although there are no pre-determined sets of criteria that can be
applied in a mechanistic way to the interpretive research, this does not mean that there are no standards for judging interpretive research.

According to many researchers, ensuring rigor in interpretive studies requires various criteria by the use of which the quality and completeness of the research process can be judged and viewed (Orlikowski et al., 1991; Walsham, 1993; 1995; Klein and Myers, 1999). These researchers have focused on qualitative research and have proposed how the nature and purpose of interpretive research varies from positivist research. Currently, there are no agreed criteria for theoretically evaluating interpretive research, however there are many such criteria available for evaluating the quality of interpretive research, such as Myers (1997) and Klein and Myers (1999) who have suggested a set of criteria for the conduct and evaluation of interpretive research in information systems. This PhD research is not concerned with obeying the scientific principles of precision and replication and instead is concerned with obtaining a theory that is compatible with evidence that is rigorous, relevant and generally useful to other areas. The following paragraphs, discuss my study in this thesis in relation to the criteria proposed by Myers (1997) for evaluating interpretive research, which suggests that interpretive research can be evaluated in terms of theory and in terms of data.

Evaluating Interpretive research in relation to the theory Myers proposed that interpretive research can be evaluated in terms of its contribution to the field and whether it has been successful in developing or applying new concepts and/or theories. The theoretical focus of this study is the factors that influence the success of data analytics endeavours in creating innovation and value in large organisations, and the results of the case study analysis was used to develop a theoretical model that explains this phenomenon. I discussed the results of the case study in terms of the categories that surfaced from the grounded theory analysis process and combined them with insights from the existing body of literature and empirical findings. Myers further argues that another fundamental concern for evaluation is whether the author proposes rich insights into the human, social and organisational aspects of information systems development and application. Rich insights into how innovation can be generated through the use of data analytics in large organisations, was developed by the analysis of the data from the field study and a considerably rich understanding of the factors that influence the success of such endeavours in creating value and innovation was offered. Furthermore, using various data collection techniques enabled a rich variety of the enterprises’ viewpoints to be captured, because each of these techniques provided a different means for the enterprises to express their perceptions of data analytics usage for innovation and the materials provided consistency in the same procedures being used for each case (Yin, 1994).
The utilisation of several techniques to obtain enterprises’ viewpoints enabled me to test one source of information against other sources, which in turn assisted in improving the quality of data and stipulated a richer, contextual basis for interpreting and validating results (Kaplan and Duchon, 1988). The last main issue considered by Myers is if the study contradicts established knowledge and hence provides richer understanding. This study created a theoretical model that established a greater and more comprehensive understanding of the factors that influence the success of data analytics in creating innovative products and services in large organisations. The organisational categories of managerial, cultural, experimentation and people factors are indicated to be significant in influencing the development of innovation through data analytics. There are currently no other comprehensive theoretical models that are proposed for the use of data analytics for innovation, and the framework developed in this study enables organisations to increase the probability of developing innovation in their data analytics endeavours.

Evaluation of interpretive research in relation to data, Myers proposes that a substantial amount of data must have been collected in order to allow for the emergence of major insights. The combination of different data techniques provided a substantial quantity of data that enabled significant insights to emerge by providing an adequate range of examples of how and why organisations succeed in utilising data analytics for innovation and business/customer value creation. Furthermore, this study provided necessary citations and quotes in each of the cases used, as a critical means of ensuring internal validity of the study. The diversity of innovative enterprises’ backgrounds in this study, offered a significantly wider context and process of value creation using data analytics in large organisations, as well as rich insights into the human, social and organisational aspects of information systems adoption and use.

Moreover, Myers believed that a well-founded interpretive research should represent various viewpoints and alternative perspectives. The inclusion of various large organisations from distinct industries that have been successful in utilising data analytics for innovation, allowed me to account for multiple viewpoints and alternative perspectives. The inclusion of enterprises from different backgrounds and industries allowed me to gather distinctive viewpoints from the participating enterprises and it provided different viewpoints on the factors that influence the success of data analytics endeavours in creating innovation. Consequently, given the diversity of the backgrounds of the enterprises selected for this study, both the context and the process of data analytics usage varies significantly and therefore a broad social context is provided.
Moreover, Myers considered the issue of whether sufficient information in regards to the research method and the research process has been provided. In this study, the application of the selected research methods, that were based on interpretive paradigm and used qualitative case study and action research, was discussed and the justification for the selection of these methodologies was provided. The interpretive approach reveals that the entire field of social relations is concerned with shared meanings, interpretations, and the production and reproduction of cultural and social reality by humans, and therefore motivates investigations into how humans enact a shared social reality through understanding human behaviour from their point of view of the world. In this study, significant attention was dedicated to the research process, in terms of the philosophical underpinnings of the research method and the practicalities of conducting the study. Details of this are provided in Chapter 4 of this study, that outlines how the data was collected, where the research was conducted, how the research developed over time, and how the data was analysed.

In conclusion of his criteria, Myers proposes that the most critical question in terms of the quality of the contribution, is concerned with the significance of the findings for both researchers and practitioners. This research has several significant strengths the most noteworthy of which is that it has been explicitly designed to develop a theory for the use of data analytics for innovation in large organisations. The study adopted a strategy of methodological triangulation that generated qualitative evidence intended to facilitate the interpretation of results.

### 7.4 Contribution of research

The contribution of this study - which presents a comprehensive view of how innovative large organisations currently use data analytics for innovation and the factors that influence the success of such endeavours, and also provides a framework to enable organisations to increase the probability of developing innovation using data analytics- can be evaluated from several perspectives.

New technologies and disruptive forces make it critical for organisations to continue innovating and meeting the customers’ needs in order to stay competitive. The emergence of big data and its immense potential for providing unprecedented competitive advantage and value, has made data analytics subject to significant attention. However, many organisations today are struggling to exploit the power of data to develop innovation and significant business/customer value in the form of innovative products and services. While there are many researchers that have tackled the problems organisations face in the of use of big data and analytics from
a technological standpoint, limited research exists in this field that addresses the organisational and practical aspects of the implementation of data analytics for innovation. This research therefore is both relevant and timely, and makes significant contributions to the understanding of use of data analytics for innovation, by not only presenting the factors that influence the success of such endeavours in large organisations, but also presenting these success factors in a cohesive and end to end framework that aims to increase the probability of large organisations developing innovation using data analytics.

This study is advantageous for researchers that seek an understanding of the factors that influence the success of developing innovative products and services using data analytics in large organisations. It can also be beneficial in studies within and across organisations by researchers who are concerned about the means by which the probability of developing innovation using data analytics can be systematically improved. Furthermore, the results of this study can act as a guide to assist decision makers in effectively utilising data analytics for innovation and it can assist practitioners and researchers understand the necessary conditions and preconditions for its success.

Another significant aspect of this study is the development of the Creative Data Analytics framework based on theory-driven case studies that explains the factors that influence or inhibit the enterprises’ ability to innovate using data analytics. The framework presents cultural, managerial, structural, and people factors as important constructs that explain data analytics usage and success for innovation.

Drawing on the rich data of the innovative organisations’ experiences, the study generated a grounded understanding of the factors that influence the success of use of data analytics for innovation. This grounded theory is valid empirically "because the theory-building process is so intimately tied with evidence that the resultant theory is consistent with empirical data" (Orlikowski, 1993). Although many scholars believe that building theory from a limited number of cases is prone to researchers’ preconceptions, I argue that the iterative comparison across cases, methods, evidence, and literature that characterises such research leads to a "constant comparison of conflicting realities that tends to "unfreeze" thinking. The process has the potential to generate theory with less researcher bias than theory built from incremental studies or armchair, axiomatic deduction" (Orlikowski, 1993).

The grounded theory developed in this study provides significant content to the understanding of the factors that influence the success of data analytics efforts resulting in innovation. Such an understanding has been absent from the research and
practice discussions of data analytics use in enterprises. The approach followed in this study predominantly focused on developing such an understanding, and therefore bringing a new set of issues to the already researched topic of data analytics. The study combines grounded theory with insights available from the innovation literature, and thereby develops a more revised general theoretical framework that enables researchers and practitioners to explain the use of data analytics for innovation in large enterprises.

The aim of using the grounded theory method in this study was to build a theory in regards to the use of data analytics for innovation in large organisations, where no theory currently exists. To date there is limited knowledge and no explanatory theory about the factors that influence the success of enterprises in the use of data analytics for the development of innovative products and services. In order to ensure the originality of this work, I conducted an extensive search of the research literature in the areas of data analytics usage, adoption, and implementation, and was unable to discover any study that uses the grounded theory methodology to develop a theory about the adoption of data analytics for innovation in large enterprises. Consequently, this study is unique as it generated a theory grounded in data collected from some of the world’s most innovative companies, about this particular phenomenon. The generated theory in this study is an original contribution to the knowledge of the field of Information Systems.

The emergent theory due to the nature of grounded theory is "abstract enough and includes sufficient variation to make it applicable to a variety of contexts related to this phenomenon’ (Strauss and Corbin, 1990). This means that the generated theory is applicable to other enterprises attempting to develop innovation using data analytics, as well as other groups such as policy makers, and members of the wider IS research community.

7.5 Conclusion

The fundamental concern of this study has been to gain deep insights into the current use of data analytics for innovation in large enterprises and the factors that influence the success of such endeavours. This study, which is based on empirical data, examines the use of data analytics for innovation in real settings of organisations and has developed a theory that considers organisational, structural, cultural and people factors, which explain the successful use of data analytics for innovation. The study’s results provide significant support to previous studies in innovation and information systems literature.
This study is presented in a descriptive manner and reveals the perceptions and experiences of the enterprises in adoption and use of data analytics for innovation. Zeller (1991) argues that studies with an interpretive perspective should not simply report the data, and rather should report "scenes" that are accounts of researchers’ engagement over time with participants in their surroundings. Furthermore, Hammersley (1992) states that "an account is valid or true if it represents accurately those features of the phenomenon that it is intended to describe, explain or theorise". This study has presented the current picture of how enterprises use data analytics in practice, and the factors that influence their success in developing innovative products and services that generate significant business and customer value using data analytics. It has set the scenes and told the stories of data analytics use from the perspectives of the enterprise cases examined.

The conclusions of this study are based on the analysis of the enterprises studied and not on a population. The goal of an interpretive study is not to make generalisations from the examined enterprises, and instead the goal is to offer understanding and/or insights in regards to the adoption and use of data analytics for innovation in large organisations. As stated by Merriam (1988) a rich description of the case enables readers to make sound judgments in regards to the transferability of the research. This study presents significant contribution towards explaining and understanding the factors that affect the success of developing innovation using data analytics and the findings provide theoretical and practical insights into this phenomenon. Finally, the research reported here contributes to what is hoped will be a continuously expanding body of empirical evidence that can increase knowledge of the use of data analytics for innovation in large organisations.
Appendix 1 - Interview questions for Australia’s largest financial institutions

Introduction

Over recent years businesses within Asia Pac have identified that either integrating data analytics capability within an organisation or transforming a business’s operations to make use of data analytics has the potential to drive a significant improvement in business performance. However, past experiences suggest that in many instances there is a gap between the business value anticipated from such an integration or transformation and the actual results delivered.

EY and UTS have reviewed recent publications in this area and drawn on our experience in this field and we have identified two specific challenges businesses face when integrating or transforming around analytics:

- Creating an appropriate people structure to support the integration or transformation, including identifying the skills and roles needed, recruiting appropriately skilled staff and structuring the analytics functions within the organisation

- Understanding and quantifying the value that can be delivered from data analytics integration and transformation within an organisation

The following interview questions are centered on your (the interviewee’s) experience with the above challenges. Your responses will contribute to a research paper that UTS and EY are writing that aims to provide insight into these two challenges based on the experience of industry. Your responses will be used anonymously in the research paper.

The interview is expected to take approximately forty-five mins to one hour.
By agreeing to be interviewed you have agreed that the interview will be recorded to assist us in accurately representing your responses. This was communicated to when you were approached about this research paper.

Can you please confirm your understanding and acceptance that the interview will be recorded (Yes/No).

IF NO then do not proceed. Thank the participant for their time.

Do you give us permission to use short quotes from this interview in our research paper (Yes/No)?

Focus Interview Questions

Question 1)

Describe your role within the organisation?

Probe for the following

- Key accountabilities, what are they required to do
- Position within the organisation, who do they report to
- Team, who do they manage, what skills does their team have, what functions does their team perform
- Knowledge and awareness of the wider operations of the organisation

Question 2)

Has your organisation has gone through any form of data analytics capabilities transformation or integration? If yes please describe:

a) The nature of the transformation/integration
b) Why the transformation/integration was initiated
c) How the transformation/integration was executed
d) What results the transformation/integration has provided
Depending on the answer to the above proceeding questions should be worded using past, present or future tense i.e. What do you think will be the challenges (are yet to integrate), what are the current challenges (currently integrating) and what do you think will be the future challenges (possible integration in the future)

<table>
<thead>
<tr>
<th>Timeline: Transformation to a data driven decision making organisation →</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motivations (Why)</strong></td>
</tr>
<tr>
<td>• External pressures faced? Is data analytics the solution?</td>
</tr>
<tr>
<td>• Is the organisation aware of these pressures?</td>
</tr>
<tr>
<td>• Awareness of data analytics as the solution?</td>
</tr>
<tr>
<td>• Awareness of data analytics as an opportunity to support future growth prospects?</td>
</tr>
</tbody>
</table>

Table A 1.1: Transformation questions

**Question 3)**

What were the key challenges that your organisation faced in the people component of the integration/transformation(s)?

**Question 4)**

How did your organisation approach structuring the people component of the integration/transformation(s) so as to address these challenges?

**Question 5)**

What were the key skills required or gaps identified during the integration / transformation(s)?

**Question 6)**

Are there any lessons you have learned from the way that your organisation approached the people component of the integration/transformation(s)?

**Question 7)**

What benefits did your business expect to realise as a result of the integration/transformation(s)?

**Question 8)**
How did your business quantify the benefits of the integration/transformation(s)?

**Question 9)**

Were there any unexpected benefits from the integration/transformation(s)?

**Question 10)**

Were any of the expected benefits of the integration/transformation(s) that were not achieved? If so, what was the reason for this? With the benefit of hindsight, what could you have done differently?

**Question 11) Closing question**

If you could tell someone who was embarking on an integration/transformation what advice would you give them?

**Close**

Thank the interviewee for their time, inform them of the next steps/timeline to release date.
Appendix 2 - Interview questions for large innovative organisations in Silicon Valley

Introduction

Over recent years businesses within Asia Pac have identified that either integrating data analytics capability within an organisation or transforming a business’s operations to make use of data analytics has the potential to drive a significant improvement in innovation capabilities and business performance. However, our research suggests that in many instances there is a gap between the business value anticipated from such an integration or transformation and the actual results delivered.

The following interview questions are centered on your (the interviewee’s) experience with driving innovation through data analytics. Your responses will contribute to my PhD research on the topic of “driving innovation in large organisations through data analytics”.

The interview is expected to take approximately forty-five mins to one hour.

By agreeing to be interviewed you have agreed that the interview will be recorded to assist me in accurately representing your responses. This was communicated to when you were approached about this research.

Can you please confirm your understanding and acceptance that the interview will be recorded (Yes/No).

IF NO then do not proceed. Thank the participant for their time.

Do you give us permission to use short quotes from this interview in my PhD (Yes/No)?

Focus Interview Questions
Appendix 2

Question 1)

Describe your role within the organisation?

Probe for the following

• Key accountabilities, what are they required to do
• Position within the organisation, who do they report to
• Team, who do they manage, what skills does their team have, what functions does their team perform
• Knowledge and awareness of the wider operations of the organisation

Question 2)

Has your organisation achieved any form of innovation that resulted in high business or customer value through the use of data analytics capabilities? If yes please describe:

e) The nature of the innovation (innovative product/service/business model etc.)

f) Why the innovation project was initiated

g) How the innovation project was executed

h) What results the innovation project has provided

Depending on the answer to the above the proceeding questions should be worded using past, present or future tense i.e. What do you think will be the challenges (are yet to use data analytics for innovation), what are the current challenges (currently using data analytics for innovation) and what do you think will be the future challenges (possible use of data analytics for innovation in the future)

Question 3)

In your opinion what kind of organisational culture supports innovation through data analytics? What are the major principles that characterise your company culture?
### Motivation (Why)
- External pressures faced? Is data analytics the solution?
- Is the organisation aware of these pressures?
- Awareness of data analytics as the solution?
- Awareness of data analytics as an innovation enabler?

### Adoption/Implementation (How)
- Current organisational culture surrounding the use of data analytics for innovation?
- What are the current organisational capabilities that enable use of data analytics for innovation?

### Use/Results (What)
- What worked? What didn’t work?
- Challenges faced in the implementation process?
- Did expectations of using data analytics for innovation met reality? Why or why not?

<table>
<thead>
<tr>
<th>Question 4</th>
<th>Question 5</th>
<th>Question 6</th>
<th>Question 7</th>
<th>Question 8</th>
<th>Question 9</th>
<th>Question 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do structured processes play a major role in your company culture?</td>
<td>Does a process exist in your company to identify how innovation can be achieved through data analytics (what problems/challenges/opportunities to focus on)? How does your company structure the very beginning of the innovation process? What are the steps that your company goes through before a product or service is actually created? Length of this? People and functions involved? Decisions made or not made? Formality of decisions?</td>
<td>What kind of indicators and assessment criteria does your company use to measure the success of ideas and to support the selection of ideas? How would you define “success”?</td>
<td>According to your experience and/or considerations, what are the key success factors for innovating using data analytics?</td>
<td>In your opinion what are the success factors for increasing the probability of data analytics resulting in innovative products and services?</td>
<td>Were there any unexpected benefits from the data analytics project(s)?</td>
<td></td>
</tr>
</tbody>
</table>
Were there any of the expected benefits of the data analytics project(s), that were not achieved? If so, what was the reason for this? With the benefit of hindsight, what could you have done differently?

**Question 11) Closing question**

If you could tell someone who was embarking on creating innovative products and services using data analytics, what advice would you give them?

**Close**

Thank the interviewee for their time, inform them of the next steps/timeline to release date.
Appendix 3 - Interview questions for managers involved in the CDA projects at Commonwealth Bank of Australia

Focus Interview Questions

Question 1)

Describe your role within the organisation?

*Probe for the following*

- Key accountabilities, what are they required to do
- Position within the organisation, who do they report to
- Team, who do they manage, what skills does their team have, what functions does their team perform
- Knowledge and awareness of the wider operations of the organisation

Question 2) How many innovation projects have you been part of? How many of those projects were focused on the use of data analytics to drive innovation? How would you describe the success of these projects in terms of achieving the expected outcomes?

Question 3) In the recent data analytics project(s) completed using the CDA Framework, how if any was this project different to other data analytics projects conducted for an innovation reason?

Question 4) In your opinion how could the CDA framework affect the definition/execution of data analytics projects?
Appendix 3

**Question 5)** How do you think the CDA framework affects the front-end of the data analytics innovation process? How does it compare to what was being done previously?

**Question 6)** How do you describe the ideas and concepts that were derived from the CDA framework in this project(s) in terms of their innovativeness? How do you think these compare to what was previously being generated using other methods?

**Question 7)** How do you predict these projects will perform in terms of creating value for the business or customer?

**Question 8)** In your opinion what are the shortcomings of the CDA framework? What is missing? What should not belong to the framework?
Appendix 4 - Data design process for the transactional data generator

In our approach, similar to agent-based models, we look at individual agents and based on the particular characteristics that are assigned to each agent (such as high/low spender) we determine their spending habits. Unlike most of the synthetic data generation tools available that do not preserve complex between attribute relationships and instead simply generate attributes as though they are independent, our data generation tool generates each column of data in dependence to the previous columns. For example a customer’s income is related to their age and suburb and is generated based on those attributes.

Our data generator uses publicly available aggregated data or known parameter distributions to generate financial transaction data. In aggregated data groups of observations are replaced with summary statistics. The main goal of statistical agencies for creating aggregated data is to limit the risk of disclosure of survey respondents’ identities or sensitive attributes, but simultaneously give statistical information in regards to the population as a whole (or certain subsets of the population) for potential future legitimate statistical analyses. Creating individualised transaction records from publicly available aggregated data to our knowledge, has not been attempted. We therefore developed a method of generating micro data from aggregated data available from statistical agencies.

The output of our data generator is financial transaction history for each customer (agent) based on that particular customer’s characteristics while ensuring that the general characteristics and distributions of purchases of the customer population remains the same as the known (real) distributions and characteristics. This way we can create a transaction data set that (even though each transaction for each customer does not necessarily follow the real trends) as a whole follows realistic trends.
Our data generator will output the financial transactional history of customers of the bank in different purchase categories such as housing, rent, food and transportation. We will also take into account the way weather affects the purchasing behaviour of the customers as an example.

Types of purchases

We consider two types of purchases that customers can make:

1. Periodic purchases: these purchase decisions are normally made with a particular pattern where purchases are made normally with a distinguished time frame in between them. This time period can be daily, weekly, monthly or yearly. Examples of these kinds of purchases are:
   - Nominal: These items can have a purchase time frame of 1 day, however their probability of purchase may change on each day of the week. Examples of this are: buying lunch, eating breakfast at a café, buying a bus ticket.
   - Seasonal: these transactions are specific to certain seasons or festivities. For example: Christmas gift buying.
   - Contractual: these purchases are tied to a contract and therefore usually the probability of these purchases being made on a regular basis is higher, especially for people with high income. Examples of this are: rent/mortgage payments, insurance payments and utilities and bills. The probability of default on these payments will depend on factors such as income, marital status and age.

2. Aperiodic: These purchases do not have particular patterns to them and consumers do not make these transactions on a periodic basis. Examples of this type of transaction are unexpected medical bills or surgeries and impulse purchases.

The model and simulator

Our transaction data generator uses publicly available data. The data sets we used to create our transaction records are mentioned below, and the details are given in Appendix 4.

We obtained the following datasets from the Australian Bureau of Statistics are the following:
Appendix 4

i. A data set that contains the number of people in different age groups in each suburb of NSW.

ii. The mean income across different age groups in NSW as well as the mean and standard deviation of income for the whole NSW population.

iii. Data set containing information about the general expenditure categories and the average spend in each income quintile; from the lowest to highest income quintile as well as the average spend in each expenditure category for the whole population.

We also used a dataset from the Bureau of Meteorology that contains information about the maximum temperature of each day in NSW in 2015.

Generating Transaction Data

We have implemented our approach as an R package. The steps taken are as follows:

a) Choose the number of customers to create transaction data history: n.

b) A data frame is created where for each customer, the suburb that they are from will be determined. This is done by first finding the probability of people coming from each suburb (dividing the number of people in each suburb by the total number of people in all suburbs). Based on these probabilities a suburb is assigned to each customer.

c) Once a suburb is allocated to each customer, the age of each customer is to be determined. Based on the probability of the age of a customer that lives in a certain suburb (number of people in each age group divided by the total number of people in the particular suburb) we assign an age to each customer.

d) Once the age and suburb are defined, income of each customer needs to be determined. For this, the function uses the mean and standard deviation of income in each age-group to create log-normal distributions of income for each age-group. Then based on the age-group of each particular customer, the customer’s income is drawn from the relevant log-normal distribution.

The problem with using the ABS data for generating transactions is that they typically do not contain the standard deviation of the values. This problem occurred when we needed to use the ABS data to generate income for each customer. Consider the following data set:
Appendix 4

<table>
<thead>
<tr>
<th>Age group</th>
<th>Mean income</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 17</td>
<td>152.4</td>
</tr>
<tr>
<td>18 – 20</td>
<td>443.9</td>
</tr>
<tr>
<td>21 – 24</td>
<td>757.1</td>
</tr>
<tr>
<td>25 – 34</td>
<td>1183.3</td>
</tr>
<tr>
<td>35 – 44</td>
<td>1393.2</td>
</tr>
<tr>
<td>45 – 54</td>
<td>1389.7</td>
</tr>
<tr>
<td>55 – 64</td>
<td>1285.8</td>
</tr>
<tr>
<td>≥ 65</td>
<td>983.9</td>
</tr>
<tr>
<td>All Ages</td>
<td>1182.4</td>
</tr>
</tbody>
</table>

Table A 4.1: Age group and mean income of customer categories

The above data set contains the mean income values for the different age groups in the NSW. In addition to this, the ABS also provides the mean and standard deviation of the income for the population as a whole.

We estimate the income distribution by a log-normal distribution (Griffiths et.al, 2012). Therefore, to generate an income for a customer with a certain age, we can use a log-normal distribution with the given mean income value for that age-group. However, in addition to the mean income, we also need a standard deviation of income for each age-group to be able to adequately choose an income from the distribution for a customer that belongs to a certain age group.

To estimate the standard deviation we first determine the relationship between the variance of income for each age-group and the variance of the income for the population as a whole. We do this using the following proof:

\[ X_i = \text{age of a randomly chosen customer.} \]

\[ Y_i = \text{income in dollars of a customer with age } X_i. \]

Now suppose:

\[ Y_i|X_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2) \]

i.e. a simple linear homoscedastic (constant variance) regression model.
For this we have:

\[
Var(Y_i | X_i) = \sigma^2
\]

\[
Var(Y_i) = E\{Var(Y_i | X_i)\} + Var\{E(Y_i | X_i)\}
\]

\[
= E(\sigma^2) + Var(\beta_0 + \beta_1X_i)
\]

\[
= \sigma^2 + \beta_1^2 Var(X_i) > \sigma^2
\]

The above depicts that the standard deviation of the population, is always higher than the standard deviation of a subset of the population. Using this information we try to approximate the standard deviation of income for the different age-groups in the following way:

According to the ANOVA method of partitioning of SSTO, if we do not use any information about the factor levels, the variability of the \(Y_{ij}\) observations will be measured in terms of the total deviations of each observations i.e.:

\[
\overline{Y}_{ij} - \overline{Y}_{..} \quad \text{(Eq. 1)}
\]

where \(\overline{Y}_{..}\) is the overall mean.

when we use the information about the factor levels however, the deviations showing the uncertainty remaining in the data will be that of each observation \(Y_{ij}\) around its respective estimated factor level mean, i.e.:

\[
\overline{Y}_{ij} - \overline{Y}_{i.} \quad \text{(Eq. 2)}
\]

The difference between the deviations in the above two equations gives the difference between the estimated factor level mean and the overall mean:

\[
(\overline{Y}_{ij} - \overline{Y}_{..}) - (\overline{Y}_{ij} - \overline{Y}_{i.}) = \overline{Y}_{i.} - \overline{Y}_{..} \quad \text{(Eq. 3)}
\]

so according to the above we can formulate our method of estimating the variance in the following way:

Total variance can be calculated as follows:
Appendix 4

\[ Var_T = \frac{1}{N_T - 1} \sum_{t=1}^{\text{number of groups}} N_i \sum_j (\bar{Y}_{ij} - \bar{Y}_{..})^2 \]

therefore the total SS will be:

\[ TSS = Var_T(N_T - 1) \]

we then from equation 2 above, can deduce that the between groups SS is:

\[ BGSS = \sum_i N_i (\bar{Y}_{i.} - \bar{Y}_{ij})^2 \]

we can then find the within the group SS (from equation 3) to be:

\[ WGSS = TSS - BGSS \]

which means:

\[ BGSS = \sum_i \sum_j (\bar{Y}_{ij} - \bar{Y}_{i.})^2 \]

Now we know that:

\[ \sigma_i^2 = \frac{1}{N_i - 1} \sum_j (\bar{Y}_{ij} - \bar{Y}_{i.})^2 \]

we now assume that the \( \sigma_i = \sigma \), meaning that this variable is the same across all age-groups. We then get:

\[ \sigma^2(N_1 - 1) + \sigma^2(N_2 - 1) + \cdots + \sigma^2(N_i - 1) = WGSS \]

Therefore, from the equation above we can find the common group standard deviation.
Appendix 4

e) After adding the income for each customer, the date and weather information will be added. For this the user needs to specify the number of days that transactions need to be generated for, and based on this the dates will be added to the dataframe. Subsequently, the weather temperature for each date will be looked up from the data set obtained from the Bureau of Meteorology, and added to the dataframe.

f) In this step we make some assumptions about the patterns of expenditure behaviour of consumers based on the temperature of the day. Specifically we specify the probability of purchasing the different spend categories and how this probability of purchase changes based on weather and the day of the week. Based on these probabilities we determine from the binomial distribution, which spend categories each customer purchases each day.

g) Once it is determined which purchases are made, we need to approximate the amount of each purchase for each customer.

To do this, first we bin the income of the customers into five categories of lowest-income, second-income, third-income, fourth-income and highest-income. This is done so that the number of bins in the income is in line with the quintiles that are used for the average spend for each spend category. The income category of each customer is then added to the customers data frame.

Now we know that in the spend categories provided by ABS only the average spend in each quintile and the overall average for each category is given. To approximate expenditure amounts, in addition to the average (mean) we need an approximation of the variance. To do this we first make the following assumptions:

1. The spend categories follow a log-normal distribution with a mean value that is equivalent to the overall expenditure mean.

2. Household income and household expenditure have perfect correlation (i.e. if family A has higher income than family B, they also have higher expenditure.

The function written to perform this task works as follows:

i. For a particular spend category, the function uses the total mean and uses a guess of the value of the variance. Using this mean and variance the function then estimates the boundaries of each quintile using the probability quantile function of the log-normal distribution:
b1<- qlnorm(0.2, mu, sig) sets the boundary of the first quintile.
b2<- qlnorm(0.4, mu, sig) sets the boundary of the second quintile.
b3<- qlnorm(0.6, mu, sig) sets the boundary of the third quintile.
b4<- qlnorm(0.8, mu, sig) sets the boundary of the fourth quintile.

ii. Once the boundaries of each quintile is approximated, we try to estimate the mean value within each quintile. We do this using the partial expectation formula:

\[ g(k) = \int_k^\infty x \ln N(x) \, dx \]  
(Eq. 4)

where \( \ln N(x) \) is the probability density function of \( X \).

The bounds of this integral will be the boundaries of each quintile (repeated 5 times to account for each quintile) and therefore using this, we are able to find the mean value within each quintile.

Once we have estimated the mean values within each quintile, we compare these values to actual mean values within quintiles that are given by ABS. The variance that gives the closest results to the ABS values will be chosen as the variance of the particular spend category. To make this comparison automatic we initially used the method of least squares method. The least squares method finds its optimum when the sum, \( S \), of squared residuals

\[ S = \sum_{i=1}^{n} r_i^2 \]  
(Eq. 5)

Is a minimum. A residual is defined as the difference between the actual value of the dependant variable and the value predicted by the model.

\[ r_i = y_i - f(x_i, \beta) \]  
(Eq. 6)

The method of least squares however is not very robust and by squaring the residuals gives more weight to large residuals. For this reason we chose the method of least
absolute deviations due to its robustness compared to the least square method and it’s resistance to outliers in the data (it gives equal emphasis to all observations).

The least absolute deviations method aims to minimise the following:

\[ S = \sum_{i=1}^{n} |y_i - f(x_i)| \]  
(Eq. 7)

Once the SD of each spend category is approximated, we are able to estimate how much each person spends in each of their transactions. In this function based on the income category of the customer, the mean expenditure value corresponding to the same quintile as income category will be chosen. Based on this mean value and the variance (we assume all quintiles have the same variance), an expenditure amount is drawn from the lognormal distribution for each transaction of each customer.
References


References


Bogdan, R.C., and Biklen, S.K., 1982, Qualitative research for education: An introduction to theory and methods, Allyn and Bacon,*Inc Boston.


Britos-Cavagnaro, L. 2013. Design Thinking Action Lab. The Context of the Innovator: Space [video file], Stanford University, Available at: https://novoed.com/designthinking/lecture_videos/615


Chaiklin, S. and Lave, J. (Eds), *Understanding Practice: Perspectives on Activity and Context*. Cambridge University Press, pp.64–103


References


References


References


References


Kaya, O. 2012, Development of an Electronic Lean Planning System for Product Development (PULSE).


References


Kumar, V. 2013. *101 design methods*.


Lee, A.S., Baskerville, R.L., Liebenau, J., and Myers, M.D., 1995, judging qualitative research in information systems: criteria for accepting and rejecting manuscripts,


Maurya, A. 2012, Running Lean: Iterate from Plan A to a Plan That Works, "O'Reilly".


Probert, D.R. 1997, Developing a make or buy strategy for manufacturing business, 1. IET.


Tamayo, Juan A., Romero, José E., Gamero, Javier, Martínez-Román, Juan A., 2015. Do innovation and cooperation influence SMEs’ competitiveness? Evidence


Torres-Padilla, Alejandro, et al, 2015. A HIGH IMPACT FRAMEWORK FOR ACCELERATING INNOVATION


References


West, M. A. 1990. The social psychology of innovation in groups.


Zuboff, S., 1988, In the age of smart machine: *The future of work and power*. Basic books.