Data-Driven Service Systems:
Applications of Social Media Dialogue Mining

A thesis submitted to The University of Manchester for the degree
of Doctor of Philosophy
in the Faculty of Humanities

2017

Szu-Yao Chien

Alliance Manchester Business School

Management Science and Marketing

Contents

Word Count: 61,768

List of Figures

List of Tables

Abstract

This thesis introduces the concept "data-driven service systems" (DDSS) to clarify the increasing adoption of big data and analytics for changing and improving the service systems where a configuration of entities (e.g. companies, customers, competitors, and diverse stakeholders) performs service exchanges and resource integration to co-create value. Grounded in a value co-creation perspective (e.g. Vargo and Lusch 2004, 2008; Grönroos 2008, 2011), this thesis proposes a DDSS framework. The framework can serve as a set of theoretical constructs, a strategic tool, and an IT prototype for researchers and practitioners to specify, design, implement, and evaluate the use of data in promoting positive system transformation.

This thesis includes three prepared journal manuscripts representing three applications that evaluate the utility of the DDSS framework. The first and second paper were conducted to examining dialogue data collected from Twitter customer care platform for addressing the research issue of service recovery. The first paper aims to provide a better understanding of customer complaint management in the dynamic service environment and uncover important activities and contexts that influence customer satisfaction. The second paper theorised a data analytical model for mining service recovery dialogues and advanced the dialogue analysis that used to be done by qualitative data analysis methods. The dialogue-mining model facilitates text mining and process mining to investigate three dimensions of dialogue including linguistic and semantic, process and relationship, and thus allows researchers to capture more insightful knowledge from a vast volume of dialogue data. Finally, Twitter dialogue data was further applied to investigate the issue of corporate social innovation in the third paper. This paper demonstrated a data-driven approach to extract and internalise stakeholder knowledge embedded in dialogues and thus, indicate opportunities for social innovation.

For each paper, novel data analytical approaches were developed to analyse unstructured data that comprises 95% of big data. In particular, text mining was used to automate information extraction in unstructured dialogue data based on specific domain knowledge (e.g. ontologies, dictionaries). In this way, text-mining approaches can provide contributions beyond the methodological and shed light into focal research domains.

Declaration

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

Copyright Statement

Acknowledgement

I would like to express my sincere appreciation to my supervisor, Dr. Babis Theodoulidis, for the opportunity to be his research student. His guidance and support throughout my PhD research have been indispensable. Without his careful supervision, this thesis would never have taken shape. Also, I gratefully acknowledge the support of my co-supervisor Prof. Jamie Burton, who constantly provided insightful advice and immense wisdom on this thesis. It has been an exceptional journey and a pleasure to work with and learn from my supervisors.

My thanks also go out to my friend Emiao Lu whom I enjoyed discussing problems with and sharing ideas during the writing-up period and also to my colleagues, Dr. Azar Shahgholianm and Carolina Martinez Troncoso, who gave me support and motivation constantly.

Especially, I would love to give big thanks to Chun-Chieh Lee who was always there to share my joys and sorrows, encourage me and give me any support I needed.

Finally, I would like to express my deepest gratitude to my parents for their sacrifice in life to financially support my PhD study and their unconditional love throughout my entire life. And, special thanks are given to my sister Szu-Ming, who always took care of my families and me, and to my grandmother, who was and will always be my life-coach.

Preface

The provided three journal manuscripts in Chapters 3–5 have been widely shared in the international marketing research conferences and presented as joint work with my supervisors. At this stage, the three papers are ready for submission, and the targeted publication outlet are Journal of Service Research, Journal of Service Management, and Computers in Human Behavior.

Chapter 3 was presented in three conferences under the titles: "*Extracting Customer Intelligence by Social Media Dialog Mining: An Ontological Approach for Customer Experience Analysis"* at AMA Summer Marketing Academic Conference, 2016, Atlanta, U.S.A. "*Tweet A Service Story: Social Media as A Service Recovery Encounter"* at AMA SERVSIG International Service Research Conference, 2014, Thessaloniki, Greece, and *"Value Co-creation on Social Media: Understanding Customer Complaint and Company's Recovery Strategy by Twitter Mining"* at AMA 23rd Frontiers in Services Conference, 2014, Miami, USA.

Chapter 4 was presented in as a conference paper under the title: "*Understanding Customer Journeys via Social Media Dialog: A Service Recovery Case Study"* at AMA SERVSIG International Service Research Conference, 2016, Maastricht, Netherland.

Chapter 5 was presented in two conferences under the titles: "*Examining Twitter Dialogs to Support Corporate Social Innovation: A Text Mining Approach*" at Global Innovation and Knowledge Academy (GIKA 2017), 2017, Lisbon, Portugal, and "*Social Media as A Co-Innovation Hub: Understanding the Idea Generation Network in the Data Rich Environment*" at 25th Frontiers in Service Conference, 2016, Bergen, Norway.

Chapter 1 Introduction

This chapter explains the motives and reasons for conducting research on data-driven service systems (DDSS). The relevant research background is discussed to highlight the multi-disciplinary nature of DDSS and gaps in the prior research. On the basis of the discussion, the research aims/objectives and a set of research questions are clarified. Finally, an overview of each chapter is provided.

1.1 Background and Motivation

In recent years, a data-driven economy has been witnessed in several industrial sectors, from healthcare, food security, energy efficiency, to intelligent transport systems and smart cities (European Commission 2014). The data-driven economy is geared by big data technologies which represent "a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery, and/or analysis" (IDC 2011, p. 6). To understand big data, several definitions have been given by practitioners, such as Gartner's "3 Vs" including volume, velocity and, variety, IBM's "4 Vs" including volume, velocity, variety, and veracity (later extended to "5 Vs" with the feature "value" added), and Microsoft's "6 Vs" including volume, velocity, variety, veracity, variability, and visibility (Wu et al. 2016).

It is worth noting that although big data is the catalyst, the data-driven transformation happens in the data market, where digital products/services are the enablers of promoting innovation, improving efficiency and effectiveness of production processes and offering a better understanding of human behaviour

(Nadkarni and Vesset 2015). A big data landscape is codified in Feinleib (2014, p. 16) to investigate two unique segments of the data market: infrastructures and applications (see Figure 1.1). The big data infrastructure market is divided into four sub-segments, including analytics infrastructure, operational infrastructure, infrastructure-as-a-service and structured database. Big data applications include vertical applications, consumer applications, business intelligence, operational intelligence, analytics and visualisation, Data-as-a-Service, and ad/media.

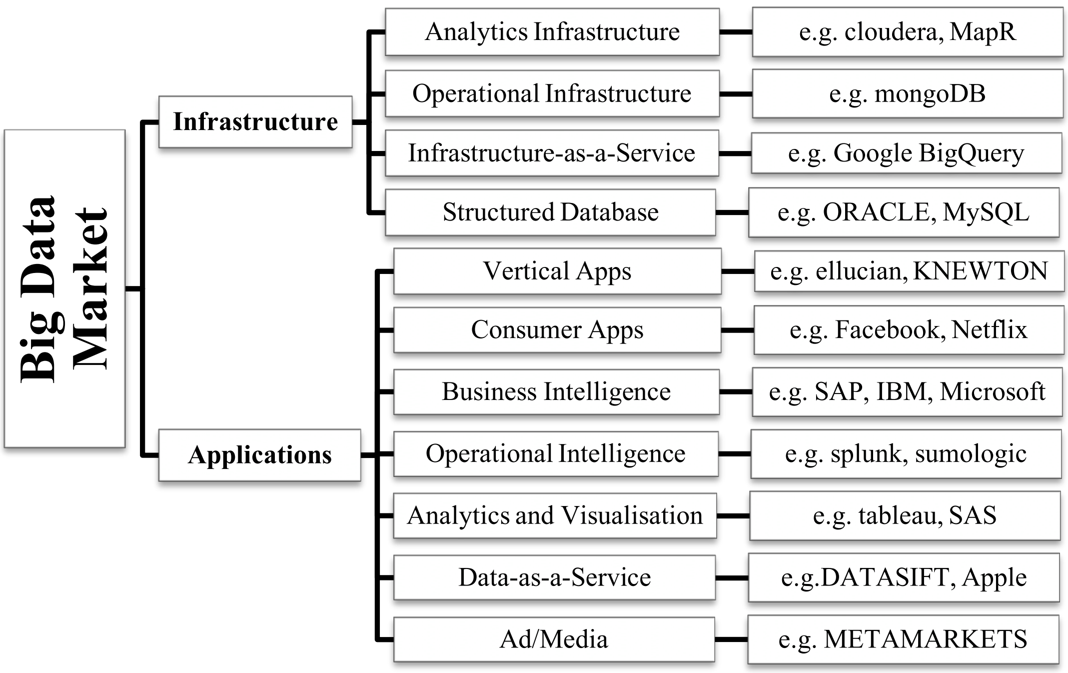

Figure 1.1 The Big Data Landscape (Feinleib 2014, p. 16)

This thesis is positioned as a data application study in the big data landscape, with a specific research interest in data analytics and business intelligence. Big data analytics (BDA) is an emerging type of business intelligence & analytics (BI&A) and closely relates to text analytics, web analytics, network analytics and mobile analytics (Chen et al. 2012). Increasingly, companies employ BDA as a solution to conduct real-time data processing and market environment monitoring (Feinleib

2014). BDA has become a vital capability for companies to achieve competitive advantage (Wedel and Kannan 2016). An empirical finding in McAfee and Brynjolfsson (2012) indicated that companies embracing a data-driven approach tend to outperform those relying on traditional decision-making methods in terms of financial and operational outcomes. In the same vein, a survey conducted on analytics-leading innovation unveiled that 61% of companies agreed that data analytics can enhance their capabilities to innovate (Kiron et al. 2012).

Big data is of limited value if companies cannot distill value from the data. "Value" is identified as the most important feature in the 5 Vs framework of big data and it is critical for companies to make better business decisions (Xie et al. 2016). Notably, the value of big data is not purely generated by BDA, but rather by the improved business practice and the enhanced relationships with focal stakeholders in the business environment. To add to the understanding of the value dimension of big data, this research adopts a data-driven service systems (DDSS) perspective, exploring how big data and BDA can be integrated into organisations' operation processes to enhance the well-being of the service systems that ogranisations, customers, and other stakeholders comprise.

1.1.1 What Are Data-Driven Service Systems?

1.1.1.1 Big Data as Cooperative Assets

In general, big data includes the following sources of data: public data (e.g. transportation, energy consumption), private data (e.g. consumer transactions), data exhaust or data by-product (e.g. online search records), community data (e.g. social media content), and self-quantification data (e.g. wristbands, mobile phone apps) (George et al. 2014). The diverse sources of big data are important resources that can later be transformed into cooperative assets potentially benefiting the cooperative actors (Xie et al. 2016).

The transformation from big data to actors' benefits highly relies on big data analytics (BDA). Grounded in the knowledge-based perspective and dynamic capacity theory, BDA is viewed by Côrte-Real et al. (2016) as a core competence that increases companies' knowledge assets and, thus, enhances companies' performance and competitive advantage. Erevelles et al. (2016) drew on a resource-based view, adding that big data can generate value which cannot be achieved by competitors when companies improve other organisational resources for utilising big data, such as physical capital (e.g. BDA), human capital (e.g. data scientists) and organisational capital (e.g. organisational structure, business process).

This thesis argues that both knowledge-based (Côrte-Real et al. 2016) and resource-based perspectives (Erevelles et al. 2016) only focus on the organisational capacities of facilitating data to create economic value and neglect the potential of big data in promoting value for other actors. To understand big data and BDA, a value co-creation view offers a better understanding (e.g. Xie et al. 2016; Kunz et al. 2017). This is because big data cannot be generated without interaction and collaboration amongst multiple people or amongst people and machines (Kumar et al. 2013). For example, customer-centric data is generated on platforms offered by or accessible to companies, including social media dialogues, website browsing and online transaction (Xie et al. 2016). Therefore, value from big data should be potentially beneficial to all data co-creators. Specifically, beyond enhancing company value, big data should also improve customer value in terms of customer satisfaction and customer experience (Kunz et al. 2017).

1.1.1.2 A Data-driven Approach for Value Co-creation

This thesis embraces a data-driven service systems (DDSS) perspective to examine how value is generated and enhanced through the application of big data. While information technology (IT) fosters more frequent interactions amongst actors, it compels companies to be more customer-centric and more service-centric (Huang and Rust 2013). The interaction between the company and other parties are the locus of value creation and value extraction, and during such a interaction, the meaning of value and the process of value creation are continuously redefined (Prahalad and Ramaswamy 2004). Direct interactions between two or more parties, such as dialogues, create a platform for value co-creation (Grönroos and Voima 2013).

A service system is the basic unit to understand value and value creation, in which individuals or networks of individuals exist, adapt and evolve (Grönroos 2008; Grönroos and Voima 2013; Vargo and Lusch 2004, 2008, 2014; Vargo et al. 2008). In the digital era, complex actor interactions within service systems are captured as big data on which companies apply their knowledge and skills to extract insights (Vargo and Lusch 2017). For example, companies provide multiple touchpoints to interact with customers and help them in contacting, purchasing, and using companies' offerings as a means to promote value co-creation. Meanwhile, data generated on the touchpoints allows companies to obtain a better understanding of the service systems and improve value co-creation.

Figure 1.2 presents the research scope of DDSS. As shown, two types of the potential value of big data are specified: information value and decision value. Wedel and Kannan (2016) stressed that information value derives from the growth of data sources in terms of volume, variety, and velocity, while decision value is gained from the improved data analytics. To understand the transformation from information value to decision value, this thesis investigates the process from data acquisition and data exploitation to data-driven decision-making.

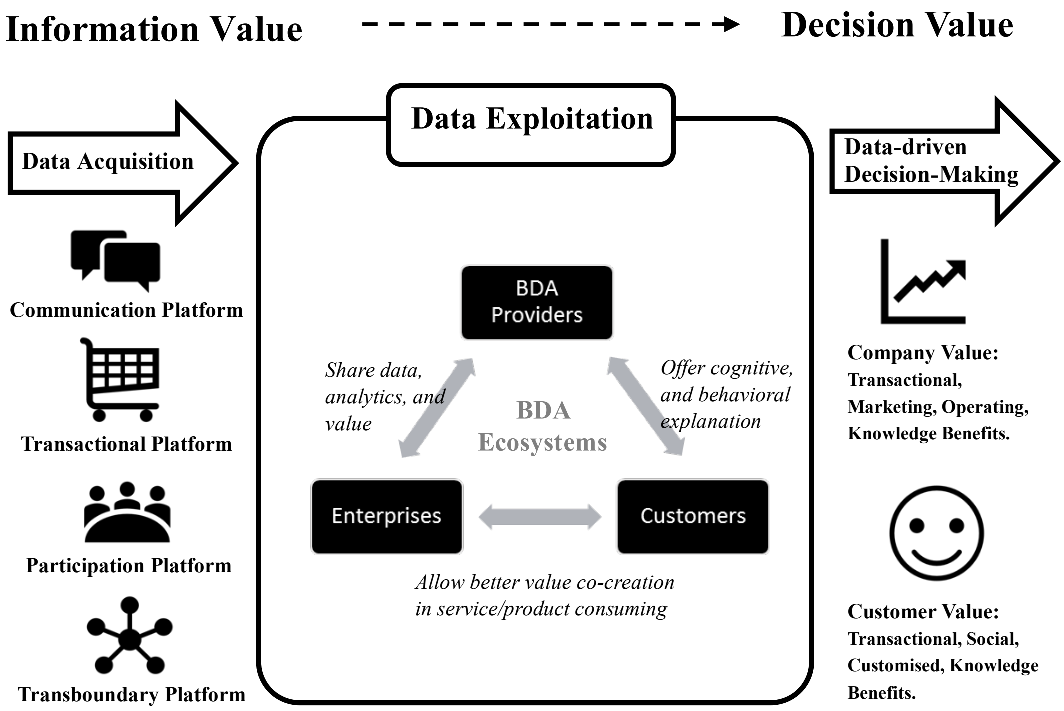

Figure 1.2 The Research Scope of Data-Driven Service Systems

The data acquisition stage seeks to obtain information value from big data within which actor value co-creation is promoted by platforms. Figure 1.2 shows four types of platforms used by companies as value co-creation enablers: a *transactional platform* supporting customers purchasing, a *communication platform* supporting both business-to-customer (B2C) and customer-to-customer (C2C) dialogues, a *participation platform* used by companies to attract customer participation in product improvement or design, and a *transboundary platform* allowing companies to obtain new knowledge from other actors in the service systems (Xie et al. 2016). Obtaining representative big data that portrays the various value co-creation mechanisms within service systems is the major task of the stage of data acquisition.

At the stage of data exploitation, BDA plays a vital role in transforming information value from different sources of big data into decision value (Wedel and Kannan 2016). Decision value can be viewed as the co-created value between companies and BDA (e.g. BDA vendors, data scientists). In contrast to traditional data analytics, BDA such as Software-as-a-Service (SaaS) acts as a resource integrator that connects two or more service systems (e.g. clients and their customers), re-bundles resources, and fosters value co-creation amongst the systems (Lusch et al. 2010). SaaS (e.g. salesforce.com) functions by collecting and analysing user-generated data on specific platforms and offering service on demand to clients through a remote web server (Lusch et al. 2010). Similarly, Data-as-a-Service (DaaS), and Analytics-as-a-Service (AaaS) later emerged as new analytics ecosystems, transforming the existing business model into a collaborative, co-developing and value co-creation model, with shared data, shared analytics and shared value (Chen et al. 2011).

Finally, at the stage of data-driven decision-making, decision value derived from BDA should help companies to achieve their short-term and long-term goals in the competitive business environment (Wedel and Kannan 2016). This thesis advocates that big data is an enabler of improving the well-being of service systems. The data-driven decision should improve company value regarding transactional, marketing, operating, and knowledge benefits and, at the same time, enhance customer value in terms of transactional, social, customised, and knowledge benefits (Xie et al. 2016). The mutually beneficial outcomes can avoid the occurrence of value co-destruction. Plé and Cáceres (2010) described that value co-destruction happens when actor interactions result in a decline in the well-being of at least one party within a service system. Value co-destruction tends to harm actor benefits and increase dynamics in service systems, such as actor exit (e.g. customer churn).

1.1.2 Social Media Big Data

In studying DDSS, this thesis particularly focuses on social media big data. Social media is identified as communication platforms on which B2C and C2C interactions occur (Xie et al. 2016). The quantity and type of social media have exponentially increased in the last decade. A recent survey pointed out that more than 90% of businesses use social media to deliver marketing communications or events (IBM 2016). The mainstream of social media adopted by companies includes blogs (e.g. Tumblr), microblogging (e.g. Twitter), collaborative wiki (e.g. Wikipedia), professional networking sites (e.g. LinkedIn), content communities (e.g. YouTube) and social networking sites (e.g. Facebook, Instagram) (Kaplan and Haenlein 2010). User-generated content on social media, such as posts, comments, pictures, video clips, or social tags, opens up new avenues to understand customers (Chen et al. 2012).

To deal with social media big data, more and more companies employ social media analytics (SMA) to capture, monitor, reply to, analyse and apply the data to enhance business practices (Zeng et al. 2010). SMA services enable marketers to listen to the "voice of the customer" and identify customers' latent needs (Hofer-Shall 2010). US computer maker Dell has even launched a social media listening command centre to monitor brand-relevant topics on social media. Mainstream SMA techniques used by practitioners to examine social media data include content mining, structure mining and usage mining (Hiroshi 2015). Content mining utilises text mining and sentiment analysis to extract users' positive/negative attitudes towards specific topics in a collection of textual data, and this technique is useful for complaint and compliment classification and customer need discovery (e.g. He et al. 2015; Ordenes et al. 2017). Structure mining facilitates social network analysis to investigate user influence within online communities (e.g. Katona et al. 2011). Usage mining focuses on user access history (clicks, likings, views) to examine marketing effectiveness and performance (e.g. Moro et al. 2016).

Prior research has made significant contributions to examining user-generated content on social media. Of the rich user-generated content, dialogue-based data tends to be under-explored, but it is crucial for understanding value co-creation amongst actors. Social media provides equal access to the companies and the customers to exchange information, and such access and transparency are critical to value co-creation (Prahalad and Ramaswamy 2004). Dialogue data consists of interrelated messages exchanged amongst multiple actors, and within the messages, information regarding service exchanges and resource integration of actors is embedded. Importantly, social media dialogue data can be easily accessed and acquired by researchers. Therefore, it is considered as a suitable data source to investigate DDSS. In this thesis, three applications using social media dialogue data to address field problems are provided.

1.2 Research Objectives and Questions

This thesis is conducted to examine the "value" dimension of big data. A data-driven service systems (DDSS) perspective is proposed to interpret value and value co-creation in the process of data acquisition, data exploitation, and data-driven decision-making. The main aim of the thesis is to improve the understanding of how the data-driven approach can enhance the well-being of service systems and create value for system actors.

Although research on big data and BDA is less theory-driven and mainly data-driven, it is important to incorporate the data-driven approaches with existing theories (Huang and Rust 2013). DDSS is built on a cross-disciplinary foundation, adopting theories, concepts, and constructs from the area of service science, marketing, and information systems (or data science). The DDSS perspective attempts to offer insights into these research fields and make contributions to integrating big data and BDA with the relevant theories.

Through this thesis, the following research objectives were achieved:

i. Grounded in multiple disciplines, a framework is developed to offer a concrete definition and a set of concepts for data-driven service systems (DDSS).
ii. The mechanisms of value and value co-creation within DDSS are explored by investigating the complex service exchanges and resource integration amongst system actors and amongst different service systems.
iii. The DDSS framework is tested in different contexts to understand its feasibility and generalisability as a conceptual framework, a management toolkit, and a prototype of data analytics.
iv. The DDSS framework is applied to address field problems using real-world datasets and provide a data analytical approach for investigating big data generated during actor value co-creation processes (e.g. actor dialogue, actor usage).

In order to achieve the aforementioned research objectives, four research questions were defined as follows:

1. What are the underlying components of DDSS? How can these components relate to the mechanisms of value co-creation within service systems?
2. How can the DDSS framework demonstrate its utility, from data acquisition and data exploitation to data-driven decision-making?
3. How can the DDSS framework be used to advance current analytical models and frameworks?
4. Can the proposed DDSS framework be applied to address field problems and offer a high applicability in different research contexts?

1.3 Structure of Thesis

This thesis is structured into six chapters. **Chapter 2** presents the research methodology based on the design science research. This chapter gives an in-depth discussion regarding the research environment and knowledge base of DDSS. Importantly, an artefact, the DDSS framework, is developed in this chapter as a problem-solving method and as a linkage between the research questions and the proposed solutions. The DDSS framework is tested and validated by three applications provided in **Chapter 3**, **Chapter 4**, and **Chapter 5**. These three applications are provided in journal manuscript format, as this is expected in an alternative thesis.

In **Chapter 3** and **Chapter 4**, two applications are conducted to investigate service recovery issues using social media dialogue data between companies and complainers. Chapter 3 focuses on customer recovery, examining the influence of dynamic factors such as the involvement of competitors and other users on customer post-recovery satisfaction. The crucial company recovery activities on recovering customer satisfaction are also analysed. Chapter 4 further examines the company's process recovery by developing a dialogue-mining framework. The dialogue-mining framework is designed to investigate the linguistic and semantic dimensions, the process dimension, and the relationship dimension of dialogues and assesses service recovery performance.

Chapter 5 provides the third application, which analyses the co-creation of corporate social innovation (CSI) embedded in social media dialogue data. Specifically, this study suggests using data to identify and reduce the cognitive distance between a company and its stakeholders and to drive CSI. Five propositions are offered based on the findings of the research, and they can serve as an operational guidance to help managers interested in implementing data-driven CSI.

Chapter 6 evaluates the proposed DDSS framework by analysing the usefulness and feasibility of the three applications. Also, this chapter discusses how this thesis addresses the research questions, summarises the thesis and indicates the research contributions, limitations and directions for further work.

Chapter 2 Research Methodology

This chapter describes the methodology – design science research – used to address the research questions. It begins by discussing the definition and process of design science research, then justifying the position, research environment and knowledge base of data-driven service systems (DDSS). Based on the justifications, a framework with key components of DDSS is then introduced. The application method is also discussed to explain how the DDSS framework will be validated using social media dialogue data.

2.1 Design Science Research

2.1.1 Definition

Design science research (DSR) is a well-developed research methodology in the information systems (IS) domain and has recently been conducted in areas such as management research (van Aken 2005) and service research (Teixeira et al. 2016). Further definition was offered by van Aken (2005), who differentiated between design science and explanatory science, stressing that explanatory science, such as natural science and sociology, is a body of knowledge offering description, explanation, and prediction for existing phenomena in the world. In contrast, design science is the knowledge that provides solutions to field problems, also known as solution-orientated knowledge (van Aken 2005). Therefore, design science research is also called "science of the artificial" (Simon 1996).

Vaishnavi and Kuechler (2015, p. 11) described design science as "knowledge in the form of constructs, techniques and methods, models, well-developed theory for performing this mapping – the know-how for creating artifacts that satisfy given sets of functional requirements". Hevner et al. (2004) stressed that design science aims to solve complex and ill-defined problems by introducing new artefacts. The artefacts in the IS domain include decision support systems, modelling tools, governance strategies, and new methods for evaluation (Gregor and Hevner 2013). The artefacts can also be new constructs, new frameworks, or new applications in business research (Teixeira et al. 2016). In short, to design is to create something that did not exist in the past (Vaishnavi and Kuechler 2015, p. 10).

Gregor and Hevner (2013) highlighted the role of artefacts in theorisation, suggesting that artefacts could be treated as a theory as they demonstrate some degrees of abstraction, which contains principles or rules and can be applied in different contexts. For instance, Teixeira et al.'s (2016) MINDS framework for service design was constructed following DSR, and the framework was used as a theorised artefact, possessing a high level of applicability in cross-industrial contexts. Another example is the widely adopted Business Model Canvas proposed by Osterwalder (2004). The Business Model Canvas is derived from a business model ontology and serves as the foundational principles for business conceptualisations and IT prototypes. The components within the business model ontology have become a popular strategic management tool in business practice. On the other hand, design artefacts are more often present in the form of applications that demonstrate a low level of abstraction and weak generalisability to other situations. Beloglazov et al. (2015), for example, designed an artefact of IT service delivery simulation models as an improvement to the old 'product line' building models. Their work is domain-specific and has a low applicability in diverse contexts.

2.1.2 Steps and Outputs of the Design Science Research Method

Previous research on DSR gave clear definitions, boundaries, and suggestions about theorisation of design artefacts. March and Smith (1995) outlined a DSR framework that contains two axes: research activities and research outputs (see Figure 2.1). Their framework specifies four types of research outputs: (1) a construct, the concept depicting field problems, (2) a model, a set of propositions that explain the associations amongst constructs, (3) a method, a set of steps guiding the conduction of tasks, and (4) an instantiation, the implementation of a design artefact in its domain. The research outputs are hierarchically layered, and the higher layers of the design product contain lower layers of research outputs (March and Smith 1995). More specifically, the instantiation operationalises constructs, models, and methods. Similarly, research activities include four types of activities: (1) build, the development of artefacts for problem-solving, (2) evaluate, the assessment of artefact performance following defined criteria, (3) theorise, the explanation of effects of design artefacts, and (4) justify, the re-examination of theories about the artefacts (March and Smith 1995). Such a four-by-four framework serves as a toolkit enabling researchers to position the DSR project in specific cells and clarify its direction (Osterwalder 2004, p. 5).

Research Activities

Research Outputs	**Build**	**Evaluate**	**Theorise**	**Justify**
Construct				
Model				
Method				
Instantiation				

Figure 2.1 The DSR Framework (March and Smith 1995)

Previous research also clarified the steps and guidelines to allow researchers to carry out high-quality DSR projects. A three-cycle view of DSR was proposed in Hevner et al. (2004) and Hevner (2007) to describe the reasoning activities in DSR projects, including the relevance cycle, design cycle, and rigour cycle (see Figure 2.2). The *relevance cycle* builds up the connection between the studied environment and DSR activities, and it begins with the identification of relevant problems or opportunities in the application environment where people, organisations, and technical systems exist. In this cycle, the artefact used for improving the environment is developed and applied to field testing. The *rigour cycle* bridges the DSR activities and knowledge base of a research project. The knowledge base contains past knowledge of the research project, such as theories and methods, experiences and expertise, and artefacts and processes in the application domain. The outputs of this cycle serve as an extension of the original knowledge (e.g., theories, methods).

Finally, the *design cycle* is at the heart of the DSR project, depicting an iterative process between building and evaluating the design artefact. This cycle lies between the relevance and rigour cycles, detecting problems and opportunities in the relevance cycle and comparing and evaluating design outputs with past knowledge in the rigour cycle. Multiple iterations may take place in the design cycle until the desirable artefacts are achieved.

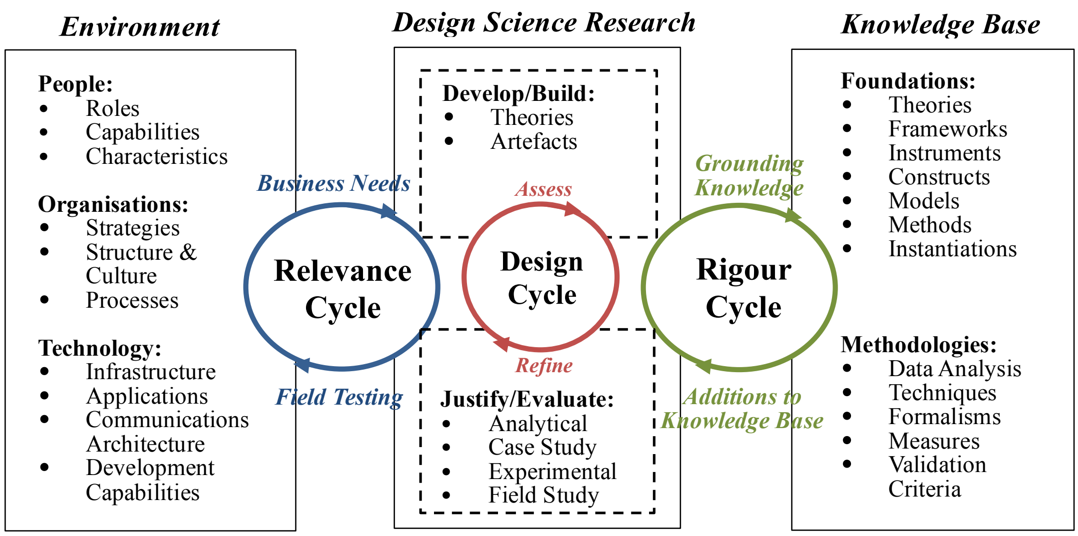

Figure 2.2 The Three DSR Cycles (Hevner 2007)

2.1.3 Justification of the Chosen Methodology

DSR is considered a suitable approach to addressing the issues regarding DDSS. This thesis sees DDSS as a “wicked problem” which is unstructured and contains confusing information and conflicting values amongst actors (Weber and Khademian 2008). Specifically, this research examines social media big data generated by multiple users who have an impact on each other in order to change the service systems. DSR is expressly concerned with such wicked problems. According to Hevner et al. (2004), DSR is conducted to address the problems featured as “unstable requirements and constraints based on ill-defined environmental contexts; complex interactions amongst subcomponents of the problem and its solution; inherent flexibility to change design processes as well as design artifacts (i.e., malleable processes and artifacts); a critical dependence upon human cognitive abilities (e.g., creativity) to produce effective solutions; a critical dependence upon human social abilities (e.g., teamwork) to produce effective solutions” (p. 81).

Hevner (2007) stressed that DSR anchors in a pragmatic philosophy that considers practical consequences and real effects as major components of both meaning and truth. However, this thesis argues that a pragmatic perspective justifying DSR based only on its solution orientation possesses a limitation in its ability to interpret and understand wicked problems. This is because pragmatism seeks practical consequences and utility that only contextually exist in the changing environment. Therefore, a critical realism perspective is taken as the philosophy stand in this thesis. Critical realism assumes that reality exists independent of human beings' cognition (van de Ven 2007). The solutions to wicked problems cannot be judged as true or false but only good or bad; more specifically, there is always more than one possible solution (Buchanan 1992).

Furthermore, critical realism also stands out as a more potent philosophical perspective compared to positivism when it comes to conducting a research methodology. Critical realism is grounded in an open social structure more similar to real social phenomena. In contrast to positivism's closed systems that offer explanations to rules or regularities based on reduced causal variations, open systems view causality as insufficient and unnecessary for understanding truth and meaning (Tsang and Kwan 1999).

Table 2.1 offers a comparison of closed and open systems in terms of their philosophical stand, mechanism, relationship, research nature, research purposes and research method. As shown, research on open systems attempts to understand the dynamic reality and complex relationships amongst entities. Such a perspective fits the purpose of DDSS research – that is, to improve the understanding of value co-creation within service systems where frequent interactions amongst actors lead to system dynamics. Moreover, big data research is, by its nature, uncontrollable since, on the one hand, the data is not created and collected based on a researcher's planned approach (e.g., experiments and surveys) but on user-generated content, and on the other hand, the data analytics may offer conflicting outcomes depending on the research purposes and statistics rules. On the basis of these arguments, critical realism is regarded as a proper underpinning of research philosophy.

Table.2.1 Comparison between Closed Systems and Open Systems (Tsang and Kwan 1999)

	Closed System	**Open System**
Philosophy	Positivism	Critical realism
Mechanism	Static reality	Dynamic reality
Relationship of Entities	One-to-one	Network
Research Nature	Controllable	Uncontrollable
Research Purposes	Prediction	Explanation
Research Method	Experiment	Observation

2.2 Position, Research Environment, and Knowledge Base of DDSS

This section discusses the position, research environment, and knowledge base of the proposed DDSS framework following the DSR method stated earlier in **Section 2.1**.

2.2.1 Position of the DDSS Framework

Figure 2.3 shows how the DDSS framework is positioned based on March and Smith's (1995) design science framework. March and Smith (1995) presented a four-by-four framework that includes 16 cells depicting practical research efforts based on the intersection of different types of research activities and research outputs. A DSR research project can cover multiple cells. As shown in Figure 2.3, the DDSS framework is positioned in the *theorise* column that explains why and how certain effects happen and how the observed behaviour can be integrated into a viable theory (March and Smith 1995).

Research Activities

Research Outputs

	Build	**Evaluate**	**Theorise**	**Justify**
Construct			Find basic components of DDSS (**Sections 2.2** and **2.3**)	
Model			Define a framework that expresses the inter-relationships amongst the components (**Section 2.3**)	
Method			Link the DDSS components to the data analytics techniques (**Section 2.4**)	
Instantiation			Apply the framework to real-world cases (**Chapters 3–5**)	

Figure 2.3 The Position of the DDSS Framework Based on March and Smith (1995)

In developing the DDSS framework, the research activities undertaken involve theorising constructs, models, methods, and instantiations. The activity in the theorise–construct cell seeks to develop the main components of DDSS, which are the basic unit of explaining big data issues. The constructs are subsequently integrated into an overarching DDSS framework (theorise–model cell), clarifying the interrelationships and rules amongst the components. Regarding the theorise–method cell, the focal point is to tie the data analytics (e.g., social media analytics) with the DDSS framework as a means to justify how and why the DDSS framework can work in practice. Finally, in the theorise–instantiation cell, a generalised theory is validated by applying the DDSS framework to address real-world problems.

In the following sections, the activities of theorising constructs, models, and methods will be discussed. In terms of theorising instantiations, it is demonstrated through three research papers provided in **Chapters 3–5**.

2.2.2 Research Environment of the DDSS Framework

Figure 2.4 shows the justification of the research environment and knowledge base of the DDSS framework using Hevner's (2007) three cycles of DSR. A brief introduction of the research environment has been given in **Chapter 1**, outlining the big data landscape and data market. In this section, the research environment is discussed based on the investigation of people, organisations, and technologies in the DDSS setting, which was earlier identified as a wicked problem. Such justification can help define the research boundary.

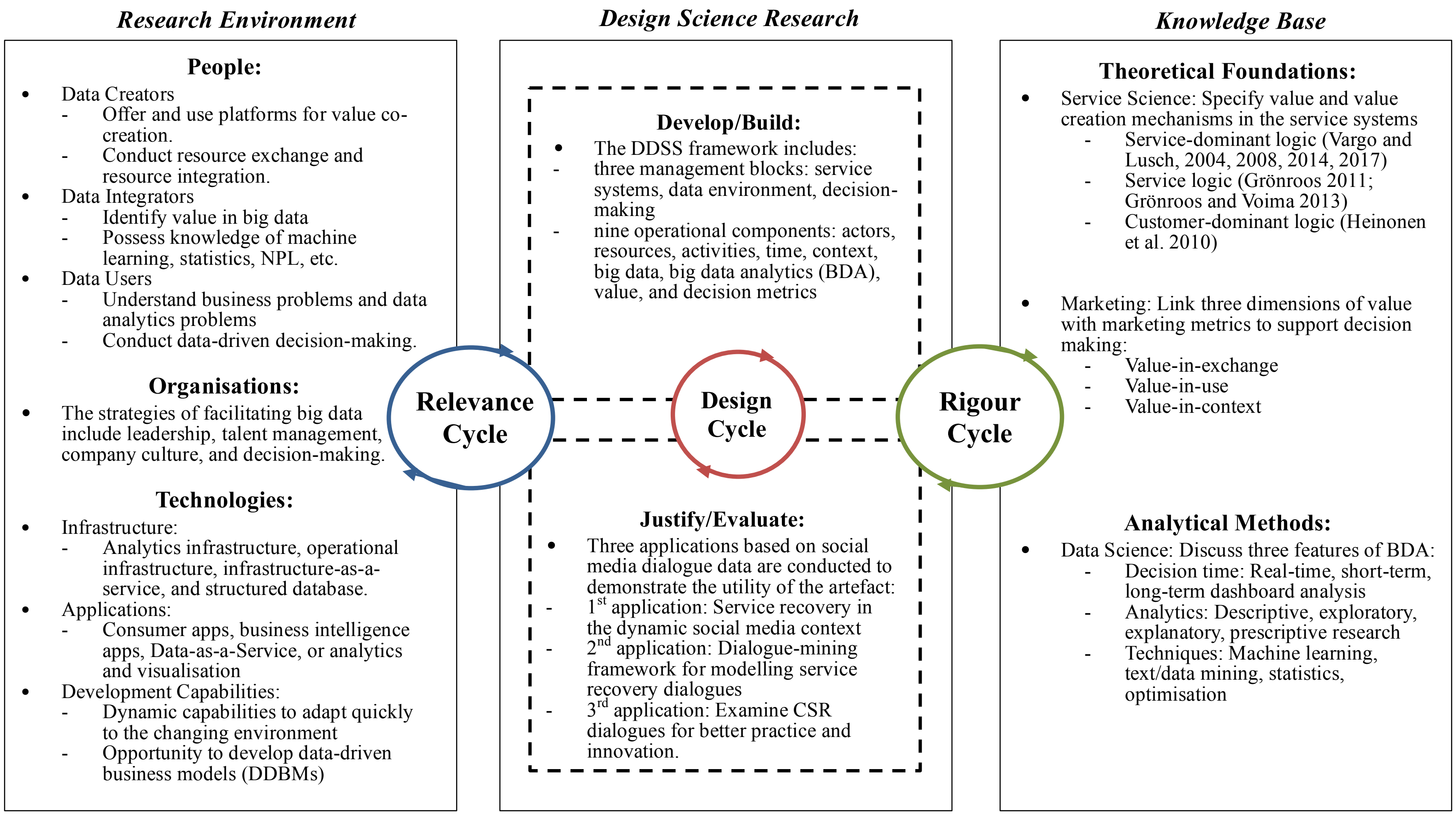

Figure 2.4 Justification of the DDSS Framework Based on Hevner's (2007) DSR Cycle

People

People in the environment can be defined by their roles, capabilities and contextual-based characteristics (Hevner et al. 2004). This thesis divides people within DDSS into the role of *data creators* whose personal details, opinions and behaviours are captured in the form of big data, *data integrators* who collect, analyse, and present the insights of big data, and *data users* who take advantage of data-driven insights to improve the well-being of a service system. People's roles within DDSS may overlap or change depending on the research contexts and business purposes.

Big data is a resource that can be generated through specific customer behaviours on company provided platforms, and this resource links to particular types of value benefitting both the company and customers (Xie et al. 2016). In the same vein, Kunz et al. (2017) stressed that big data is a crucial enabler for companies to optimise value generation in all customer engagement touchpoints. Therefore, data creators need to possess capabilities such as providing platforms and using platforms to co-create value through resource exchange (e.g., offering personal information) and resource integration (e.g., using a wearable device to track the body's activities).

Big data is of limited value if it is not transformable into actionable insights. This task falls to data scientists, who are recognised as data integrators within DDSS. Companies have identified the need to nurture and retain data scientists as a business strategy for seizing opportunities enabled by big data and analytics (Fosso Wamba et al. 2015). Named as the "sexiest job of the 21st century", data scientists are in high demand in industries such as information technology, marketing, finance, and government (Davenport and Patil 2012). Data integrators do not directly participate in the value co-creation process, but they identify value embedded in data and deliver it to data users in the form of actionable insights. The key capabilities of a data scientist include mathematics, machine learning, artificial intelligence, statistics, natural language processing, database management, and optimisation (Carillo 2017).

Data users make efforts to change DDSS to allow for better value co-creation. Decision makers are important data users who rely on data-driven value to enhance a company's performance, revenue and reputation and, simultaneously, provide a more personalised service and build a deeper relationship with customers (Kumar et al. 2013). Carillo (2017) highlighted that analytics capabilities should not only pertain to data scientists but also be possessed by all employees, especially decision makers and managers. Data users need to have knowledge of both business issues (e.g., IT-enabling industry revolutions) and data analysis issues so that they can benefit from data-driven decisions.

Organisations

All the aforementioned roles (data creator, data integrator, and data user) are involved in or connected to organisations. Hevner et al. (2004) stressed that people perceive business needs in the environment, and business needs can be evaluated based on an organisation's strategies, structure and culture, and business processes. The main issues of organisations within DDSS are how they can facilitate big data to improve products/services and develop dynamic capabilities to adapt quickly to the changing business environment (Xie et al. 2016). To manage changes effectively, organisations should develop five management strategies regarding big data: leadership, talent management, company culture, decision-making, and technology (McAfee and Brynjolfsson 2012).

Organisations within DDSS should be able to deal with both business issues and big data issues. McAfee and Brynjolfsson (2012) stated that a leadership team is key to a company's success in the big data era, and the team should be constructed to tackle business issues, such as detecting opportunities, understanding the market, articulating clear goals, and dealing with customers, employees, business partners, and other stakeholders. To cope with big data issues, organisations should also have a data analytics team with data scientists and other professionals involved to solve problems in a way that cannot be achieved by traditional statistics methods (McAfee and Brynjolfsson 2012).

In addition to the organisational structure tailored to reap big data benefits, organisations within DDSS need to embrace the culture of utilising data to change and improve business processes. This relies on an organisation's strategy of decision-making with respect to obtaining and using the right data, distilling value from data and applying the insights to business processes (McAfee and Brynjolfsson 2012). More importantly, empirical findings in Lavalle et al. (2011) indicated that big data and analytics are not the main challenges; it is the organisational culture that does not support information sharing and the decision-making process where organisations lack understanding of using big data analytics to improve practice. Similarly, Kumar et al. (2013) emphasised that the key challenge of managing decision-making is that managers tend to rely on their own evaluation rather than data-driven insights if the data-driven value is not clear.

Technologies

Technologies within DDSS can be assessed based on infrastructures, applications, and development capabilities (Hevner et al. 2004). The term big data previously simply referred to data of vast volume that could not be processed by traditional database systems. It later obtained widely recognised definitions such as 3 Vs (volume, variety, and velocity) and 5 Vs (volume, variety, velocity, veracity, and value). One of the key big data technologies is the most widely adopted open source software platforms, Apache Hadoop, which is derived from the MapReduce framework and implemented in the Java programming language (Chen et al. 2012; Tambe 2014). Apache Hadoop has several sub-projects, including Cassandra, Pig, Hive, and the Hadoop Distributed File System for handling different tasks of the Hadoop cluster interface, communication, and processing flow (Tambe 2014).

To implement the big data software and data environment on computer clusters, big data infrastructures are also highly developed. Feinleib (2014, p. 16) classified the big data infrastructures into operational infrastructure, Infrastructure-as-a-Service, structured databases, and analytics infrastructure. These big data infrastructures are designed to enable data tasks, including data acquisition from multiple sources, data transformation, data repositories, running data through high-performance analytic engines, and reporting and visualisation (Boinepelli 2015). Big data infrastructures were traditionally costly, requiring local data warehouses and installing and maintaining complex software. The advance of cloud computing makes Infrastructure-as-a-Service (IaaS) available (e.g., Amazon Elastic Compute Cloud) and therefore reduces the costs of the up-front investments in storage and computing infrastructure (Feinleib 2014, p. 86).

The advance of big data hardware and software has led to dramatic industry revolutions and allows companies to offer more and better personalised services, build deeper service relationships, create more profitable customers and shift to more service-centric business models (Rust and Huang 2014). When it comes to the applications of big data, one of the most important business models is consumer applications. Netflix, as an example of consumer applications, uses a Hadoop-based infrastructure to analyse customers' viewing habits, giving customised content recommendations (Harris 2012). Other types of big data applications include business intelligence applications (e.g., SAP), Data-as-a-Service, (e.g., DATASIFT), and analytics and visualisation (e.g., SAS) (Feinleib 2014, p. 16). Big data applications are often designed to suit multiple business purposes. Social media such as Facebook, Twitter, YouTube, and LinkedIn are categorised as consumer applications, but they also offer Data-as-a-Service by trading user-related data to other applications providers and offer data analytics to allow users to track their social performance through dashboards.

In terms of development capabilities, in addition to the continuously evolving big data infrastructures and applications, the research issue regarding data-driven business models (DDBMs) also recently caught researchers' attention. Hartmann et al. (2016) stressed that DDBMs facilitate big data as the key resource to running a business. In their empirical research, six types of DDBMs were identified: free data collectors and aggregators, analytics-as-a-service, data generation and analysis, free data knowledge discovery, data-aggregation-as-a-service, and multi-source data mash-up and analysis (Hartmann et al. 2016). Importantly, big data and analytics are viewed as an enabler of business model innovation (Zolnowski et al. 2016). Prior research conducted on analytics-leading innovation unveiled that analytics can improve organisations' capabilities to innovate, and analytical innovators are found to use more data, manage information more efficiently and are speedy in processing and analysing data (Kiron et al. 2012).

Discussion on the research environment of DDSS indicates the potential problem settings where this research can make contributions. As stated earlier, DDSS is established on multi-disciplinary foundations, and thereby the outcomes of this thesis can provide new solutions such as a management toolkit or a prototype of data analytics to understand and manage the focal research areas.

2.2.3 Theoretical Background of the DDSS Framework

In this section, the theoretical foundations of DDSS are discussed to provide a linkage between the knowledge base and the research environment. By doing this, the critical concepts, components, and mechanisms of the DDSS framework (design artefact) can be captured.

As discussed in **Section 2.2.1**, the purpose of the DDSS framework is to theorise constructs, models, methods, and instantiations (March and Smith 1995). It requires a wide review of the foundations of DDSS regarding the theories, frameworks, instruments, models, and methods in prior research (Hevner et al. 2004). Big data is, by its nature, grounded in a multi-disciplinary perspective. Extant literature sees big data as "the next management revolution" (McAfee and Brynjolfsson 2012), the "next frontier for innovation, competition, and productivity" (Manyika et al. 2011), and as "...opportunities allowed by the information revolution" (Goes 2014, p. iii). Big data and analytics create a disruptive impact on business strategies and business models for nearly all sectors (Hartmann et al. 2016; Weill and Woerner 2015). In developing the DDSS framework, this thesis draws on service science to provide the fundamental knowledge for dynamic digital ecosystems, and then the relevant research areas such as marketing and data science are discussed.

Service Science

Companies embracing a data-driven approach are found to shift from a product-centric to service-centric business model (Zolnowski et al. 2016). The e-book, as an example, derives from traditional paper books and shows a digital transformation from searching, acquiring, paying, and subscribing (Weill and Woerner 2015). Tesla, the automobile maker, is another case of data-driven transformation. Tesla analyses the data collected from customers' cars, identifies when the cars are due for repairs and notifies the customers automatically (Carillo 2017). Digital technology and the big data-driven transformation lead today's companies to operate in a value ecosystem within which they seek to know more about end customers and partner closely with other users – even competitors (Weill and Woerner 2015). Specifically, enterprises arising from big data analytics adapt to or create system dynamics by connecting different parties within the service ecosystems, capturing the actor-centric data and offering computational tools (Vargo and Lusch 2017).

In developing the DDSS framework, it is necessary to understand the core concepts – value and value (co-)creation. A changing view of value has been highlighted in the early research, from a company-oriented to a customer-oriented perspective. Normann and Ramírez (1993) suggested a concept of value constellation, emphasising that business strategies should be positioned on reconfiguring roles and relationships amongst constellation actors (e.g. suppliers, partners, customers) to promote value creation (Normann and Ramírez 1993). Prahalad and Ramaswamy (2000, 2004) stated that the market is transforming from a place for trading to a place for conversation and interactions, and thus, enables customers to co-create value and tailor unique experiences (Prahalad and

Ramaswamy 2004). The concept of value and value co-creation becomes the backbone of service science. To understand value and value co-creation in the service systems, three main streams of service science are discussed, namely service-dominant logic (SDL), service logic (SL) and customer-dominant logic (CDL).

SDL, as a paradigm shift from goods-dominant logic, was proposed by Vargo and Lusch (2004). Goods used to be viewed as value carriers for which customers are willing to pay monetary resources in exchange. Vargo and Lusch (2004) argued that the process of exchange should be viewed as a service, and goods are the means of service provision. Services are the exercise of competences, such as knowledge and skills, and allow one party to benefit another – "value-in-exchange" (Vargo and Lusch 2004, 2008; Lusch and Vargo 2014). One of the most important contributions of SDL is that it challenges the traditional value creation in goods-dominant logic, clarifying that value is always co-created. In other words, companies cannot deliver value, and value is always co-created when customers consume the value propositions provided by companies – "value-in-use" (Vargo and Lusch 2004; Vargo et al. 2008).

Following SDL, Maglio and Spohrer (2008, p. 18) defined service systems as "a configuration of people, technologies, organisations and shared information, able to create and deliver value to providers, users, and other interested entities through service". The service systems proposed by Maglio and Spohrer (2008) drew on the static environment, where the influences amongst multiple service systems were not taken into consideration. Lusch and Vargo (2014) extended the concept of service systems and proposed service ecosystems. They emphasised the dynamic nature, stressing that a service ecosystem is "a relatively self-contained, self-adjusting system of resource-integrating actors connected by shared institutional logics and mutual value creation through service exchange" (Lusch and Vargo 2014, p. 161).

Figure 2.5 shows a visualisation of service ecosystems proposed by Lusch and Vargo (2014). In the service ecosystems, multiple networks of actors (e.g., company–supplier network, company–customer network) are involved, and amongst the networks are complex resource exchanges (e.g., products, money, skills) and value co-creation through resource integration (e.g., using products, learning new knowledge). Notably, every actor within the service ecosystems is a resource integrator (Vargo et al. 2008). All actors – not just companies – can develop and offer value propositions of their service to enable value co-creation (service-to-service exchange) (Lusch and Vargo 2014). For example, customers can offer services in the way of their opinions, ideas and monetary resources that allow companies to co-create value and better business practice.

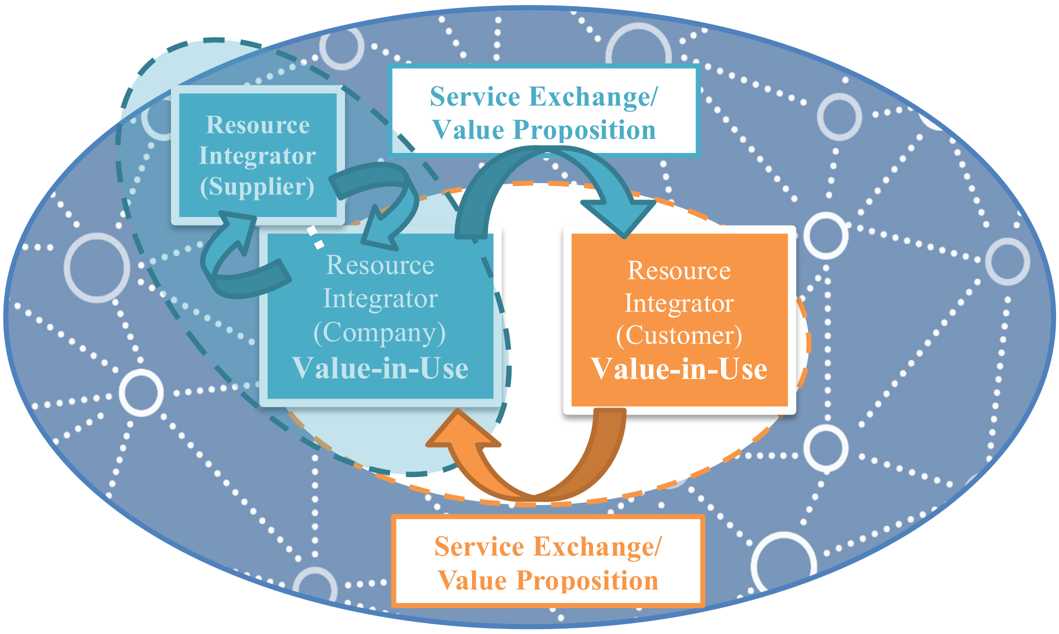

Figure 2.5 Service (Eco)systems Based on SDL (Lusch and Vargo 2014)

Grönroos and Gummerus (2014) criticised SDL as a perspective of company-driven value creation where companies, customers, and other actors participate in the process leading to value for customers. Grönroos (2008, 2011) argued that SDL is not a one-applies-to-all solution, since the meanings of value creation/co-creation are interpreted differently from the perspectives of the service providers and the customers. Instead, Grönroos (2008, 2011) suggested a service logic (SL) perspective centred on customer-driven value creation. SL clarifies that value creation is customers' creation of value-in-use, and a company's role during such a process is as a value facilitator engaging with customers' processes to enable reciprocal value creation (Grönroos 2011; Grönroos and Voima 2013).

A model of three value-creation spheres was specified in Grönroos (2011) and Grönroos and Voima (2013): provider sphere, customer sphere, and joint sphere. Figure 2.6 demonstrates the service systems based on the three-sphere model. In the provider sphere, the company acts as a value facilitator producing potential value in the form of value-facilitating goods and services as outputting resources (Grönroos 2008b). Value facilitation is not part of value creation, as real value only occurs in the customer sphere when the potential value is later turned into value through customers' self-value generating processes (Grönroos 2008b). The customer sphere is an experiential sphere where customers apply their experiences with resources and activities in specific contexts (Grönroos and Voima 2013). In this sphere, customers can also co-create value with other people (Grönroos and Gummerus 2014). Notably, it is only in the joint sphere that value co-creation occurs between the company and the customers through direct interaction (Grönroos and Voima 2013). The joint sphere is similar to the value co-creation of SDL, in which resource exchange and resource integration happen, and value propositions of the company are transformed into customers' value-in-use (Vargo and Lusch 2004, 2008).

If value can be co-created, it may be destroyed throughout the customer's value-creating process (Grönroos and Gummerus 2014). Plé and Cáceres (2010) took an opposite view of value co-creation, defining value co-destruction as "an interactional process between service systems that results in a decline in at least one of the systems' well-being" (p. 431). Value co-destruction is one of the major reasons for service failures and negative customer experiences. This is because value is subjectively evaluated from the provider's and the customer's perspectives (Echeverri and Skalen 2011). For example, a customer who purchased a car may damage the value due to a lack of maintenance, and this type of value co-destruction results from the customer's misuse (Plé and Cáceres 2010). It is important for service providers to effectively manage provider–customer interactions in the joint sphere and improve the understanding of customer value-creation processes and contexts (Grönroos and Voima 2013).

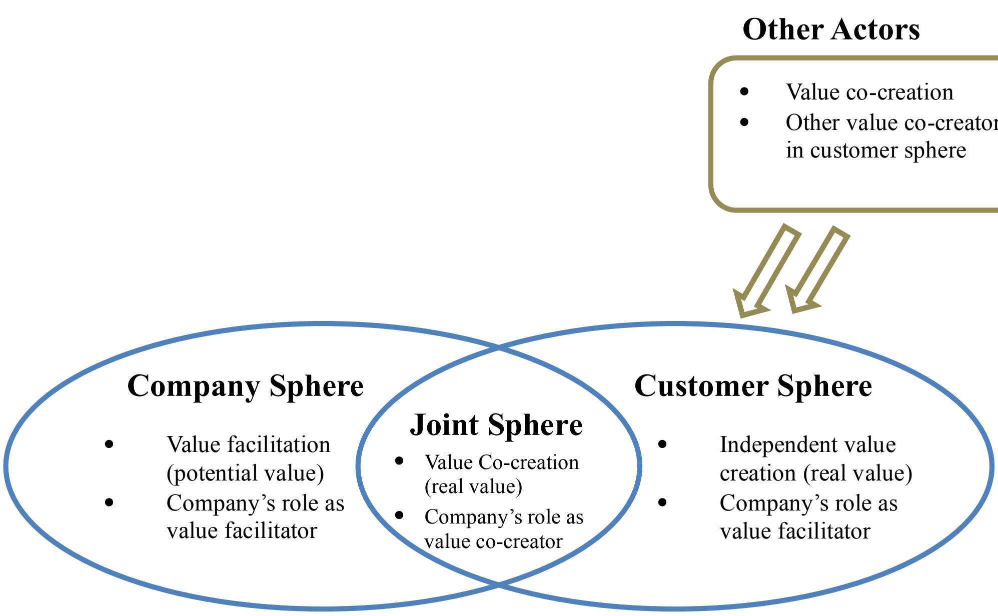

Figure 2.6 Service (Eco)systems Based on SL (Grönroos and Voima 2013)

Similar to SL, customer-dominant logic (CDL) rises as another stream of service science in reacting to SDL. CDL proposes a customer-centric perspective that focuses on the customer's constellation of activities, actors, and experiences and clarifies how providers and their offerings are embedded in the customer context (Heinonen et al. 2010). CDL highlights that customers are the priority in marketing practice rather than services and service (eco)systems (Heinonen and Strandvik 2015). More specifically, companies embracing CDL focus on customer-related aspects rather than products, services, or costs (value propositions in SDL, and potential value in SL). In CDL, value is contextual, and thus provider–customer interactions are not the critical mechanism of the value formation; rather, it is how customers use and apply service offerings in their lives and ecosystems (Heinonen and Strandvik 2015). CDL is important for companies to manage the customer experience, because instead of following companies' pre-designed service processes, customers are actively choosing, participating, consuming, and leaving at any point of the service process.

Figure 2.7 shows the service systems in the perspective of CDL (Heinonen and Strandvik 2015). Two worlds are proposed – the customer's world and the provider's world – and the overlapping area of the two worlds is the interaction arena. A temporal process through the pre-service, service, and post-service stages is highlighted in the CDL. During this process, the provider becomes involved in the customer's activities and experiences. According to Heinonen et al. (2010), customers' value-in-use is created based on their service experience embedded in their own context, and this is similar to the value-generating process in Grönroos and Gummerus (2014).

.

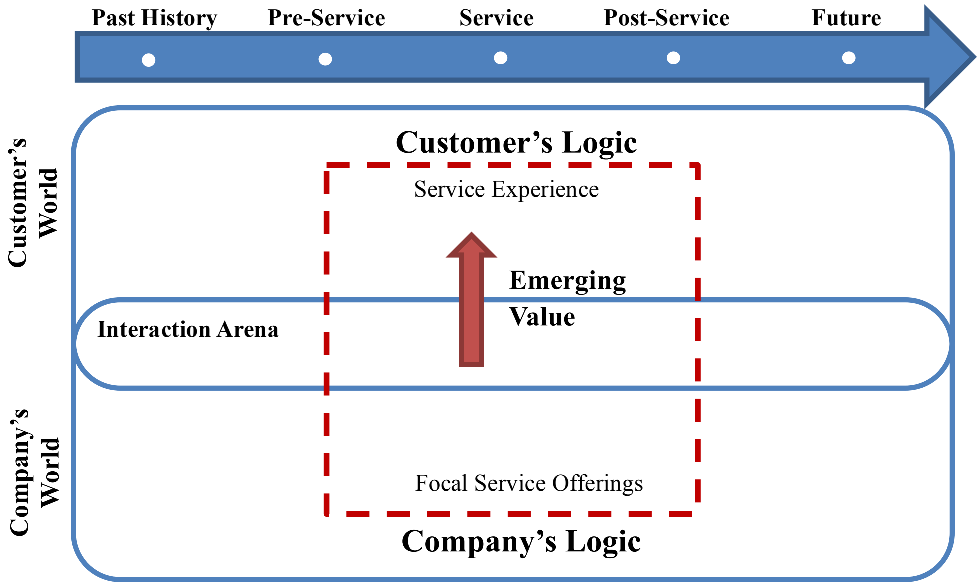

Figure 2.7 Service (Eco)systems Based on CDL (Heinonen and Strandvik 2015)

Table 2.2 provides a comparison amongst the three research streams and justifies the components of service systems mentioned in each stream. Since this thesis does not aim to create debates amongst the three streams but to understand the main mechanisms and essential elements of service systems, components possessing high agreement and high generalisability are selected for further investigation of DDSS. They are value, actors, resources, activities, time, and context.

Table 2.2 Comparison of the Three Streams of Service Science

Research Streams	Main Argument	Components of Service Systems	Authors
Service-Dominant Logic (SDL)	1. A paradigm shift from goods-dominant logic to service-dominant logic. 2. Service is the fundamental basis of exchange.	Value (value-in-exchange, value-in-use), Value co-creation, Value proposition, Interactions, Exchanges, Resources,	Vargo and Lusch (2004, 2008); Lusch and Vargo (2014); Löbler (2013); Wilden et al.

	3. Value is co-created by multiple actors and is always uniquely and phenomenologically determined by the beneficiary. 4. All social and economic actors are resource integrators. 5. Value co-creation is coordinated through actor-generated institutions and institutional arrangements. (based on the five axioms of SDL in Vargo and Lusch 2017)	Activities, Processes, Context	(2017)
Service Logic (SL)	1. A perspective of customer-centric value opposite to the company-driven value of SDL. 2. Value creation and co-creation of service are distinct. 3. The interaction process is where value emerges. 4. Companies can act as a value facilitator engaging in the customers' value creation process to enable reciprocal value creation. 5. Value is sometimes destroyed throughout the customer's value-creating process.	Value (value-in-use) Value co-creation, Value creation, Value facilitation Interactions, Actors, Activities, Resources, Experiences, Processes, Context	Gronroos (2011); Grönroos and Voima (2013); Grönroos and Gummerus (2014)
Customer-Dominant	1. A customer-dominant view contrasting with	Value (value-in-use),	Heinonen et al. (2010);

Logic (CDL)	a provider-dominant view. 2. Companies' offerings are embedded in customers' lives and ecosystems. 3. As opposed to being deliberately created, value is formed in use separately for customers, providers, or any other actor involved. 4. Value-in-use may be emotional or symbolic and reside outside interactions.	Value formation, Actors, Time, Activities, Experiences, Resources, Context	Heinonen et al. (2013); Heinonen and Strandvik (2015); Mickelsson (2013); Tynan et al. (2014)

Value and Marketing Metrics

The foundations of value have been thoroughly studied in service science. However, a missing link between the value of service systems and marketing metrics is recognised in this thesis. Nothing within SDL is viewed as a theoretical framework, SL is positioned as an analytical approach, and CDL used as a marketing management strategy specifies the connection between the value of service systems and quantifiable marketing metrics. Although the follow-up research has shed light on the specific aspect of marketing value, such as customer loyalty and repurchase intentions (e.g., Kuzgun and Asugman 2015; Leroi-Werelds et al. 2014), there is still an absence of a holistic view that links the value of service systems with marketing metrics.

It is important to establish a link between metrics and value when the field continuously advances towards a tangible performance metric that evaluates the impact of marketing (Kumar 2016), especially as practitioners still make decisions based on their subjective evaluation if they cannot “see” the benefits driven from marketing metrics (Kumar et al. 2013). Notably, linking value of service systems with marketing metrics does not mean a shift back to goods-dominant logic (e.g., monetary value, marketing ROI), but provides quantifiable criteria of value to support decision-making.

According to Farris et al. (2010, p. 1), marketing metrics are a measuring system used for quantifying trends, dynamics, and characteristics, as well as being a communicational tool for explaining phenomena and sharing findings and results of future events. Practitioners need to choose and develop suitable metrics, as marketing has multiple dimensions, including attitudinal, behavioural and financial facets, and the outcomes of marketing events may be viewed differently from various metrics (Hanssens and Pauwels 2016). For example, a digital marketing campaign may be regarded as a success in a metric of website visits but does not have an impact on sales revenue.

There are many popular marketing metrics highlighted in prior research. Kumar and Reinartz (2006) outlined three categories of marketing metrics related to customer relationship management:

(1) Traditional marketing metrics (e.g. market share, sales growth),

(2) Customer-based metrics (e.g. acquisition rate, share of wallet), and

(3) Strategic customer-based value metrics (e.g. customer lifetime value).

Mintz and Currim (2013) highlighted 10 marketing-mix decisions in addition to the associated marketing metrics:

(1) Traditional advertising (e.g. impressions, reach),

(2) Internet advertising (e.g. hits, visits, page views),

(3) Direct to consumer (e.g. number of responses),

(4) Social media (e.g. number of followers, number of tags),

(5) Price promotions (e.g. trial / repeat volume),

(6) Pricing (e.g. price premium),

(7) New product development (e.g. attitude toward products/brands),

(8) Sales force (e.g. new customer retention rate),

(9) Distribution (e.g. product category volume), and

(10) Public relations (PR) / sponsorships (e.g. volume of coverage by media).

Importantly, with more availability of big data enabled by digital technologies and platforms, new marketing metrics are introduced. Kumar et al. (2013) examined data-driven services marketing, specifying a set of marketing metrics of digital data sources, such as search queries (e.g. web traffic), social media (e.g. reach, likes, shares), blogs (e.g. valence of posts, visits, bounce rate), community forums (e.g. membership size), and incentivised referrals (e.g. acquisition rate, share rate). Hoffman and Fodor (2010) investigated 50 marketing metrics to understand the ROI of social media in promoting brand awareness (e.g. number of visits), brand engagement (e.g. number of followers), and word-of-mouth buzz (e.g. number of shares).

In this section, the SDL perspective is used to explain the relationship between marketing metrics and the value of service systems. The three dimensions of value are value-in-exchange, value-in-use, and value-in-context. As stated earlier, value-in-exchange is the application of resources that have value potential for benefitting others, and it can also be viewed as service-for-service exchanges (Vargo and Lusch 2004, 2008, 2014, 2017). Resources in value-in-exchange particularly relate to operant resources (e.g. knowledge, skills) rather than operand resources (e.g. money, goods) (Vargo and Lusch 2004, 2008, 2014, 2017).

Value-in-use refers to the real value co-created through use by end users (Vargo and Lusch 2004, 2008, 2014, 2017), or the real value emerging during end users' usage of resources (Heinonen et al. 2010; Grönroos and Voima 2013). The concept of value-in-use is extended into value-in-context, which is defined by the proposition "value is uniquely and phenomenologically determined by the beneficiary" (Vargo and Lusch 2017, p. 47). Actors are traditionally regarded as owners of resources who have control of resources within their context. Chandler and Vargo (2011) argued that each actor can define their own contexts as well as the resources within the contexts, and simultaneously, the dynamic and fluidity of the market is framing and being framed by contexts. The principle of value-in-context explains how resources become resources and specifies the contexts where resources are embedded and transformed into value-in-use through usage by actors (Chandler and Vargo 2011). Though value-in-context has been viewed as an extension of value-in-use, this thesis embraces the early definition of value-in-use built on customer-to-company interaction and value-in-context built on the networks of actors (or service systems).

From the above discussions, a cross-tabulation analysis is used to examine the associations between value dimensions and system actors' value and then to specify how the value dimensions and the actors' value can be evaluated by the relevant marketing metrics. Table 2.3 shows the cross-tabulation analysis drawn on Kumar et al.'s (2013) customer-based metrics. As shown, in the dimension of value-in-exchange, customer value may be obtained from the exchange of solutions embedded in goods. Company obtained value is sales revenue for its survival, and the relevant metrics are share of wallet or a new customer acquisition rate. As for the dimension of value-in-use, customer value may be co-created from problem-solving or convenience of use, and company co-created value is customer satisfaction and service quality review. The relevant metrics can be customer churn rate and service failure rate. Finally, the value-in-context dimension focuses on the temporal process, during which customer value (e.g. trust) and company value (e.g. reputation, customer relationship) are changing all the time. The evaluation metric may be customer lifetime value.

Table 2.4 presents another cross-tabulation analysis based on Hoffman and Fodor's (2010) social media marketing metrics. The value-in-exchange dimension includes customer value obtained from the exchange of information and knowledge and company value obtained from brand awareness. The relevant metrics include the number of visits, impressions or page views. In the value-in-use dimension, the customer value is co-created through learning, experiences, satisfaction, and fun, while company value is co-created through brand engagement evaluated by metrics such as the number of comments, likes, or shares. Finally, value-in-context focuses on the social context, where customer value emerges from self-affirmation or altruism, and company value may be word-of-mouth views that can be assessed by the metric share of voice.

With the understanding of the associations amongst value dimensions, actors' value, and marketing metrics, marketers can make more accurate marketing decisions pertaining to the value they can have an impact on and the correspondent marketing performance they can achieve.

Table 2.3 Linking Service Systems Value to Customer-based Metrics (based on Kumar et al. 2013)

Value in Service Systems	Customer Value	Company Value	Marketing Metrics
Value-in-Exchange	• Goods • Skills	• Sales revenue	• Acquisition rate • Acquistion costs • Share of wallet • Sales growth
Value-in-Use	• Problem-solving • Convenience • Satisfaction	• Customer satisfaction • Service quality review • Staff knowledge and skills	• Churn rate • Retention rate • Expected service failure and recovery rates
Value-in-Context	• Trust • Commitment	• Reputation • Customer loyalty • Customer relationship	• Customer lifetime value

Table 2.4 Linking Service Systems Value to Social Media Metrics (based on Hoffman and Fodor 2010)

Value in Service Systems	Customer Value	Company Value	Marketing Metrics
Value-in-Exchange	• Information • Knowledge	• Brand awareness	• Number of unique visits • Number of page views • Number of reviews posted • Number of members/fans • Number of installs of applications • Number of impressions • Number of bookmarks

			• Number of reviews/ratings
Value-in-Use	• Learning • Experiences • Satisfaction • Fun	• Brand engagement	• Number of comments • Number of active users • Number of “likes” on friends’ feeds • Number of user-generated items (photos, threads, replies) • Usage metrics of applications/ widgets • Impressions-to-interactions ratio • Rate of activity (how often members personalize profiles, bios, links, etc.)
Value-in-Context	• Self-affirmation • Altruism • Catharsis • Vengeance	• Word-of-mouth	• Share of voice • Frequency of appearances in timeline of friends • Number of posts on wall • Number of reposts/shares • Number of responses to friend referral invites

Data Science

The justification of service sciences and marketing metrics provides theoretical foundations of DDSS. The knowledge base of DSR also includes relevant methodologies, such as data analytics, techniques, and measures (see Figure 2.4). Big data holds the answers to many field problems, and data analytics are the key to the answers. Big data analytics (BDA) has been viewed as one of the major research streams in data science (others are big data infrastructure and transformation and impact) and also in other research disciplines, especially marketing (Goes 2014). The increasing availability of big data requires relevant BDA to extract actionable insights and to enable companies to make better decisions following tailored marketing metrics (Wedel and Kannan 2016).

As shown in Figure 2.8, using BDA to support decision-making should consider three aspects: decision time, analytics and techniques (Goes 2014). BDA regarding decision time is often supported by the use of dashboards. A dashboard is developed to satisfy the need for integrating diverse business activities in an overview of both short-term marketing performance and long-term health of the marketing assets (Ambler 2003). Analytics can support different types of research activities, such as data visualisation, exploratory research, explanatory research, and predictive research (Goes 2014). Wedel and Kannan (2016) also identified four levels of marketing analytics activities: (1) descriptive research summarises and visualises the variables and describes the tendency of a dataset in an exploratory way, (2) diagnostic explanatory research aims to test hypotheses and estimate relationships between variables, (3) predictive research is designed to forecast variables of interest and conduct simulations of the effect of marketing control settings, and (4) prescriptive research focuses on determining optimal levels of marketing control variables. Notably, many existing statistical and econometric models are not tailored to deal with big data, especially unstructured data which consists of the majority of big data, and these research activities require new types of techniques for data transformation and data modelling.

Big data modelling techniques are, in general, rooted in disciplines such as statistics, computer science and information systems. Several techniques have already been in use in marketing in reaction to the digital transformation and service revolution, including data/text mining, machine learning and fast Bayesian methods (Rust and Huang 2014). The growth of user-generated content (Twitter, Facebook, blogs, reviews) opens up new ways to understand customers and leads Business Intelligence & Analytics 1.0 (BI&A 1.0) to the era of BI&A 2.0. The data in BI&A 1.0 are mostly structured data collected by companies through various enterprise systems and stored in commercial relational database management systems (Chen et al. 2012). BI&A 2.0 is characterised by advanced modelling techniques, such as web-based unstructured content mining, information retrieval and extraction, opinion mining, social media analytics, social network analysis, and spatial-temporal analysis (Chen et al. 2012). Marketing research examining user-generated content has, for example, applied text mining to investigate market structure (Netzer et al. 2012), adopted speech act theory to examine implicit and explicit language used to express sentiment in user reviews (Ordenes et al. 2017), and used data-mining techniques to predict social media performance (Moro et al. 2016), to name a few.

Data-/text-mining techniques often facilitate machine learning methods.[1] Machine learning aims to design algorithms allowing computers to evolve behaviours based on a considerable dataset and discover embedded knowledge for automatic decision-making (Chen and Zhang 2014). BDA relies heavily on statistical machine learning with well-established mathematical models and algorithms such as

[1] Machine learning wiki: https://en.wikipedia.org/wiki/Machine_learning

Bayesian networks[2], hidden Markov models[3], and support vector machines[4] (Chen et al. 2012). These statistical techniques have been widely applied to text analytics, audio analytics, video analytics, social media analytics, and predictive analytics (Gandomi and Haider 2015), and they have become popular in marketing practice due to their outstanding predictive performance and black-box nature (Wedel and Kannan 2016).

Although machine learning techniques enable researchers to use probability models without predictor variables, academic marketing research may be reluctant to adopt such "one solution fits all" models and estimation methods (Wedel and Kannan 2016). This is because marketing discipline contains several sub-domains, such as customer relationship, branding, and marketing-mix strategies, and marketing metrics. The "one size fits all" analytics are not desirable and offer limited understanding to marketing theories.

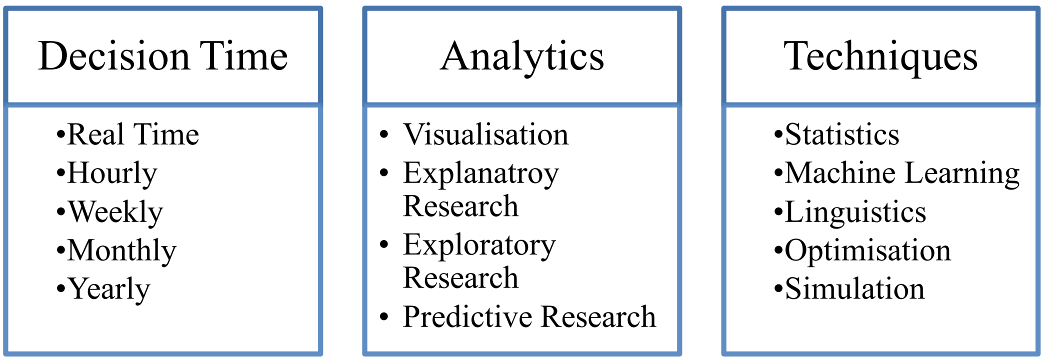

Figure 2.8 Big Data Analytics (BDA) and Decision Support Systems (Goes 2014)

[2] Bayesian networks wiki: https://en.wikipedia.org/wiki/Bayesian_network
[3] Hidden Markov models wiki: https://en.wikipedia.org/wiki/Hidden_Markov_model
[4] Support Vector Machine wiki: https://en.wikipedia.org/wiki/Support_vector_machine

2.3 The Data-Driven Service Systems (DDSS) Framework

On the basis of the justification of the research environment and knowledge base (see Figure 2.4), this section introduces the DDSS framework and explains its rationale in terms of theorising constructs and theorising models (see Figure 2.3). As for theorising methods and theorising instantiations, they will be provided in **Section 2.4** and **Chapters 3–5,** respectively.

The term data-driven service systems (DDSS) has not been defined in the published scholarly research, yet a growing number of service research and information systems research studies have been conducted to explore how big data and BDA can be applied to enhance service practice and customer value (e.g. Kunz et al. 2017; Xie et al. 2016). To add to the poorly-defined DDSS, this thesis takes an value co-creation perspective, highlighting two major goals of DDSS: first, DDSS is functioned on introducing changes into service ecosystems by constructing new value propositions; second, DDSS aims to promote better value co-creation and improve the well-being of service systems through the adoption of big data and BDA.

Built on multidisciplinary foundations, the framework makes use of a set of concepts and definitions discussed in **Section 2.2**, codifying a main structure of DDSS with three management blocks: *service systems*, *data environment*, and *decision-making*. Within the three blocks are a set of operational components: *actor*s, *resources*, *activities*, *context*, *big data*, *BDA*, *value* and *decision metrics*. Figure 2.9 presents a visualisation of the proposed DDSS framework, disclosing the positions and relationships of each component.

The *service systems* are the forum where the company interacts with customers to conduct service exchanges and value co-creation (Lusch and Vargo, 2014). Five key components engaged in value co-creation are identified:

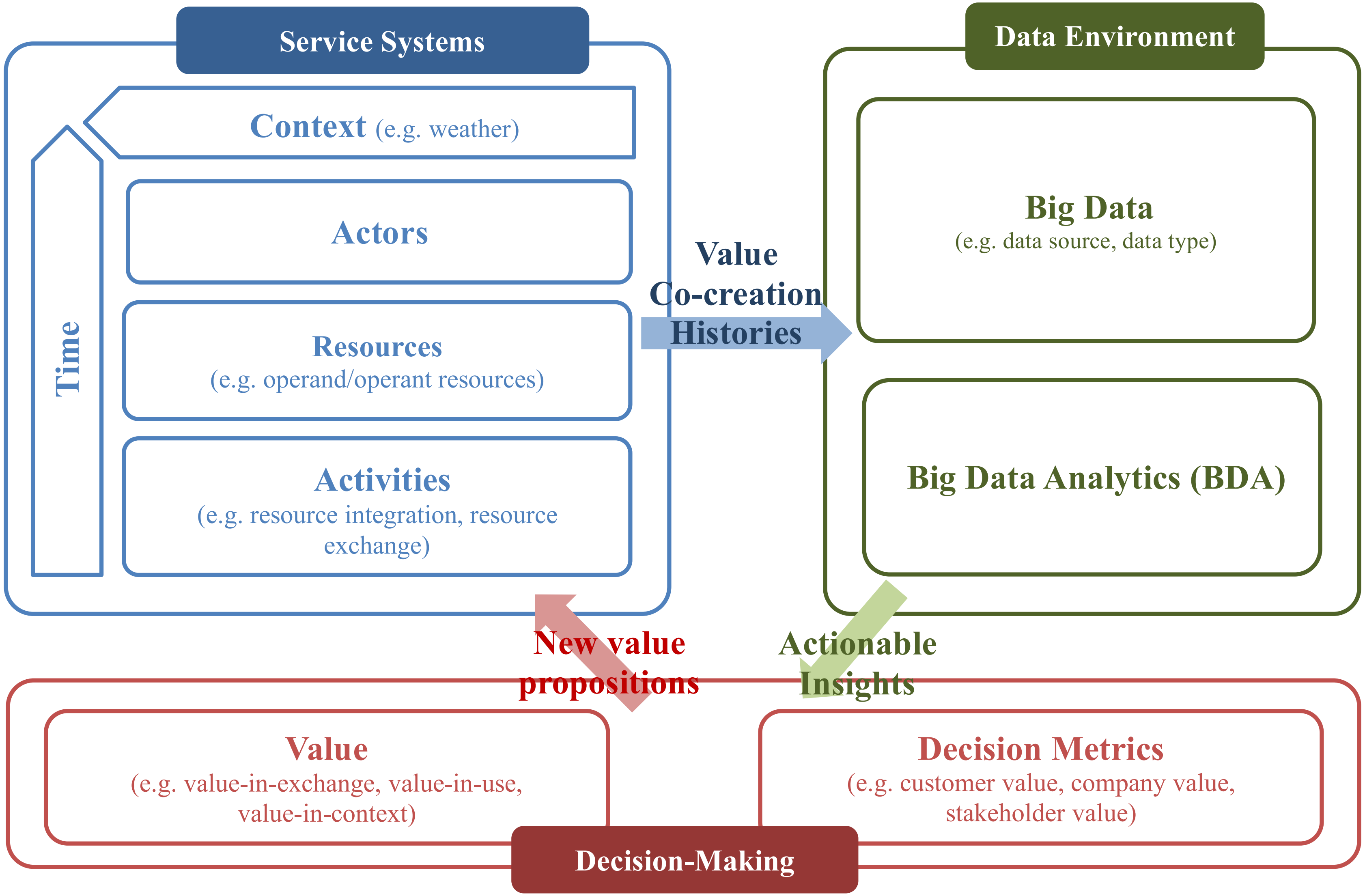

Figure 2.9 The DDSS Framework

- **Actors.** Actors are the basic unit of a service system to promote value co-creation. The scale of service systems can be defined from micro (e.g. customers and companies) to meso (e.g. communities, companies, and suppliers) and macro (e.g. the society) levels (Webster and Lusch 2013). Usually, service systems are built on networks of actors who develop relationships with others and rationally co-create value through service exchange (Lusch and Vargo 2014). The actors' role within the network is changeable depending on how resources are offered and consumed. Actor-to-actor (A2A) exchange brings the linear co-creation process of value towards a more complicated and dynamic exchange system (Lusch and Vargo 2014). The dynamic networks of actors are also known as a value constellation (Patricio et al. 2011). The actors' position and role in the network will determine their abilities for resource integration, thus influencing value co-creation (Lusch and Vargo 2014).

- **Resources.** One key characteristic of service systems is the service-for-service exchange through which actors generate, integrate, and apply resources to achieve personal goals and make others better off (Grönroos and Voima 2013; Lusch and Vargo 2014). Two types of resources are identified: operant and operand resources (Vargo and Lusch 2004, 2008; Vargo and Akaka 2012). Operant resources (often featured as intangible and dynamic resources, such as knowledge and skills) are the driver of exchange in the service systems and possess capabilities of acting on operand resources (often featured as tangible and static resources, such as goods and money) to provide benefits (Vargo and Lusch 2004). All actors, such as customers, employees, and social communities, are resource integrators and, more specifically, are operant resources that integrate resources to realise value (Vargo and Akaka 2012). This thesis distinguishes actors from other types of operant resources and sees them as a critical component of service systems. This is because, as well as acting on operand resources, actors also take actions on operant resources in ways of applying knowledge and skills of their own or others to achieve specifically goals.

- **Activities.** Service systems function based on two major activities: exchange of resources and integration of resources (Vargo and Lusch 2004, 2008). Resource exchange activities channel value propositions from one service system to others, yet value is only co-created through resource integration activities (Vargo and Lusch 2004). The concept of a value constellation is highlighted by researchers (e.g. Patricio et al. 2011) to describe more complex value co-creation activities in the dynamic networks of actors. Primarily introduced by Normann and Ramírez (1993), value constellation describes how a company's offerings can be provided jointly with other parties' offerings to create customer value. Notably, the value constellation highlighted in Lusch and Vargo (2014) points out that all actors are resource integrators. In other words, the role of actors is often changed or repositioned within the value constellation depending on the resource exchange and integration activities.

- **Time.** Service systems are not static but dynamic, as system actors can participate, exchange resources, consume resources, or exit. In both SL and CDL, the interaction between customers and companies is continuous value co-creation over time (Grönroos and Voima 2013; Heinonen et al. 2010). SDL views time as a resource that is uncontrollable by actors but is integrated by the whole service system (Vargo et al. 2008). The existence of service systems varies based on their size and time. Webster and Lusch (2013) suggested that the understanding of service systems should move from small systems, defined by factors such as the transaction time frame, to meso or macro systems, defined by the relational time perspective. In other words, the survival of small systems tends to be influenced by a single actor's actions, such as customer churn. However, the small systems are important for researchers to understand the operations of service systems because the boundary is clear and manageable.

- **Context.** SDL highlights that "value is always uniquely and phenomenologically determined by the beneficiary" (Vargo and Lusch 2017, p. 47). In the service systems, several dynamic factors, such as weather or laws (Vargo et al. 2008), influence resource exchange and resource integration activities. These dynamic factors are context-based and affect value co-creation amongst actors.

The five components describe the mechanisms of service systems, and these components formulate service events that are recorded in big data. The *service systems* block is connected with the *data environment* block (the second service system) by big data, and the connected blocks form a meso-level service system. Notably, the role of actors is often re-defined based on the nature of services they provide and receive within or across service systems. For example, in the micro system, companies play the role of data creator, but in the meso system, they act as a data integrator or data user.

The *data environment* block includes two components: *big data* and *big data analytics (BDA)*. Below, the roles of *big data* and *BDA* in the DDSS framework are clarified:

- **Big Data.** Big data cannot be created without interactions in the service systems. Kumar et al. (2013) described big data as the digital data generated during interactions between actors, such as data from social media, or human interactions with platforms, such as clickstreams. Xie et al. (2016) stressed that big data is a resource that can be transformed into cooperative assets, and such assets contain potential present or future economic benefits which can be obtained and managed by the cooperative parties through service exchanges. Three features of cooperative assets are recognised: interactive, integrative, and bilateral (Xie et al. 2016). These features follow the SDL perspective and are consistent with the transformation of big data through service exchange (interactive feature noting big data as a product of interactions), resource integration (integrative feature referring to an actor's capabilities of collecting, analysing and using big data), and value co-creation (bilateral feature describing benefits of big data offered to cooperative actors). However, Xie et al. (2016) only emphasised the micro service systems where only the company and its customers are present. Instead, the DDSS framework suggests that a meso or a macro perspective of service systems is more suitable to examine big data, as the company often relies on other network actors (e.g. academics, data scientists, BDA vendors) to jointly enhance value co-creation processes (Vargo and Lusch 2017).

- **Big Data Analytics (BDA).** In contrast to traditional data, big data relies on specific BDA to capture, process, and analyse it (McAfee and Brynjolfsson 2012). Prior research (e.g. Wedel and Kannan, 2016) pointed out that BDA acts as a crucial service that distills insights from big data. Companies' BDA capabilities often develop through service exchanges with third parties, such as BDA vendors, or data scientists. For example, companies need to provide access to enable the BDA (e.g. Analytics-as-a-Service) to collect and investigate the data. Here, BDA is viewed as an operant resource through which companies can acquire actionable insights into big data. Therefore, the capabilities of using BDA are the process of the company's value co-creation.

Big data that contains events and histories of service systems can only improve the value co-creation and well-being of micro systems when it is transformed into new value propositions. The actionable insights generated in the data environment carry potential value propositions that can be used by the management teams of companies. The meso system perspective is still applicable to explain the connection of the *data environment* block and the *decision-making* block. However, service ecosystems are dynamic and continually influenced by the changes of value propositions (e.g. new technologies, new policies), thus affecting the mechanism of value co-creation (e.g. new user habits). Therefore, a meso system may grow into a macro system in which more parties participate.

In the *decision-making* block, the components of value and decision metrics are justified as follows:

- **Value.** The insights produced in the data environment are fed back to the organisation to develop new/improved value propositions in the form of new/improved products, services or processes. The value propositions aim to enhance the value co-creation in the micro service system. Three types of value have been specified in **Section 2.2**: value-in-exchange, value-in-use, and value-in-context. To realise value, value propositions are the driver that is exchanged, co-created, and phenomenally co-shaped amongst users. The development of new value propositions should consider the benefits of the whole service system, as value can be co-created and also be co-destroyed. Value co-destruction occurs when one party's value co-creation harms the other parties' value (Plé and Cáceres 2010). An example of value co-destruction is that of a company providing a low-priced product to customers through the exploitation of labour. It is important to understand how the three types of value function for each actor or each network of actors in different contexts.

- **Decision Metrics.** The data-driven insights for developing new value propositions need to be linked to specific marketing metrics to allow for observable results through quantifiable criteria. As discussed in **Section 2.2**, the decision metrics should be based on relevant value aspects possessed by actors in the micro system. This can support the decision-making accordingly.

2.4 Applications of the DDSS Framework

This section explains how the DDSS framework can be evaluated through applications to real-world big data. The design of an artefact includes the development and evaluation stages (design cycle). Evaluation can offer feedback about the artefact, thus improving its quality. Evaluation methods include an analytical approach, experimental approach, case study or field study (Hevner et al. 2004). In this thesis, social media dialogue data is used for the evaluation of the DDSS framework in the field settings. Three applications are provided as a means of assessing the utility, quality, and efficacy of the DDSS framework. More details of the evaluation will be discussed in **Chapter 6**.

2.4.1 Social Media as a Service Ecosystem

A survey conducted in 2015 showed that 96% of businesses use social media to deliver marketing communications or events (IBM 2016). Social media is an actor networking platform and a marketing ecosystem (Hanna et al. 2011). The nature of social media is similar to the actor-to-actor (A2A) network highlighted in Lusch and Vargo (2014). Through social media, actors such as customers, companies, suppliers, competitors and other stakeholders are connected, and they conduct service exchange (e.g. word of mouth) for value co-creation (e.g. dialogue, event participation). Value co-creation within an A2A network is dynamic, as the competition for resources, as well as the cooperation of resources, occurs frequently (Lusch and Vargo 2014). For example, multiple companies simultaneously interact with customers (operant resource) to compete and to win their attention, forge a relationship, and encourage a purchase. Similarly, multiple customers can join together in social media dialogues to co-shape a unique experience involving a company’s service (Baird and Parasnis 2011). The dynamic and evolving feature of social media A2A networks satisfies the service ecosystems or value constellation scenario.

Singaraju et al. (2016) stressed that social media serve as a systems resource integrator that implements higher levels of resource configurations by interconnecting other resource integrators (e.g. customers, companies). Social media are an important intermediary, and their core service for exchange is to channel value propositions from one actor to another (Singaraju et al. 2016). In return, social media obtain the service in the form of user-generated information (big data) to co-create value. Other mechanisms of the service exchange of social media include offering user-generated data or advertising opportunities to other systems resource integrators such as companies, analytics vendors, and marketing agencies in exchange for monetary resources.

The service exchanges of social media become increasingly important for companies to obtain competitive advantages. In addition to using the network service, today’s companies increasingly facilitate social media big data to monitor markets, acquire competitive intelligence and mine customer insights (Zeng et al. 2010). Social media data is useful for understanding virtual user interactions, and this is demonstrated in both a semi-structured format (e.g. social tags, likes, and shares) and an unstructured format (e.g. social dialogues) (Khan 2015). The rich information of social media data helps companies to address specific business problems through the use of social media analytics.

2.4.2 Existing Social Media Analytics (SMA)

To examine social media big data, the use of social media analytics (SMA) has become an emerging research area (Ruhi 2014). According to Zeng et al. (2010), SMA refers to the tools, applications or frameworks that collect, monitor, analyse, and visualise big data on social media. Recently, SMA has been used to address several marketing issues. For example, Fan and Gordon (2014) applied SMA to track the online reputations of a hotel brand to analyse service problems and customer satisfaction.

In business practice, the terms social listening, social monitoring, and social intelligence are often related to SMA. Social listening is similar to social monitoring, and it is designed to examine social "buzz" (e.g. mentions, likes, sentiments) towards a product, a service or a brand on social media sites (Moe and Schweidel 2014). Social listening helps companies to understand brand health and marketing campaign performance. Social intelligence is based on social listening and provides multi-dimensional knowledge (e.g. influencers, market trends, competitors) with which practitioners can take action (Moe and Schweidel 2014). Companies increasingly use social media intelligence as a means of conducting market research and formulating marketing strategies.

Multiple SMA techniques have been developed to tackle different business issues, including demographic analysis, geographic analysis, text analysis, image analysis, and influencer identification, profiling and scoring (IBM 2016). These techniques can be applied to aid content marketing, such as investigating both paid and earned social media content on Facebook, Twitter, YouTube or Instagram (e.g. text and image analysis), and also help targeted advertising through demographic and geographic analysis (Wedel and Kannan 2016). It is important that the use of SMA techniques follows specific SMA frameworks, otherwise the social media data may only provide social noise rather than intelligence (Moe and Schweidel 2014).

Several SMA frameworks are provided in prior research. Fan and Gordon (2014) proposed a CUP framework containing three stages of SMA: capture, understand and present social media data. Stieglitz and Dang-Xuan (2013) offered an analytical framework systematically outlining approaches and the relevant techniques for tracking, monitoring and investigating social media content in the political context. Khan (2015) took a cyclic view, detailing an SMA process to achieve business objectives, including identification, extraction, cleaning, analysing, visualisation, and interpretation.

This thesis highlights a number of challenges encountered in existing SMA. First, prior SMA frameworks focusing on purely analytical processes neglect that the diverse sources and kinds of social media data require a holistic view that enables both academics and practitioners to justify the associations between user behaviours and data, between data and data analytics, and between analytics-enabling insights and business strategies. Second, SMA investigates the conversations and interactions amongst users to extract insightful social intelligence. Zeng et al. (2010) pointed out that social interactions are often context-dependent. However, SMA is usually built on generalised algorithms or lexicons for processing user-generated data in a timely manner. This affects accuracy and potentially misleads findings in the data. Zeng et al. (2010) also mentioned that SMA is multi-disciplinary, yet the level of integration between SMA (informatics) and domain sciences (e.g. social science, business research) tends to be low. In other words, research on SMA is often methodology-centric, and there is a missing link between SMA and domain knowledge.

Three applications based on social media big data and SMA will be discussed in Chapters 3–5. Through the applications, this thesis sheds light on the identified challenges of SMA. The DDSS framework serves as a solution offering an integrated view of SMA following the evaluation of three management blocks: *service systems*, *data environment*, and *decision-making*. Regarding the issue of data analytics, the applications explore the interrelated content generated from social dialogue interactions rather than user posts. This approach provides solutions to the limitation of processing context-dependent data in the SMA. Moreover, domain-specific analytical approaches are used, such as ontologies and dictionaries that theorise the methods of SMA and contribute to domain knowledge.

2.4.3 Evaluation by Social Media Dialogue Data

Of the rich social media data, dialogue data is the most important information that contains details of user interactions regarding purposes, formats (e.g. texts, pictures, videos), processes and outcomes. Social media dialogue data is considered suitable for examining the DDSS framework as it is representative for a value constellation. Interaction is at the centre of service systems, as is big data (Kumar et al. 2013). Dialogue data contains identifiable actors, actors' world views, and social contexts, thus providing visible co-creation processes and outcomes.

2.4.3.1 Data Collection

To collect social media dialogues, the researcher consulted a netnography approach. According to Kozinets (2002), netnography, or ethnography on the Internet, is "a qualitative consumer research methodology that uses the information publicly available in online forums to identify and understand the needs and decision influences of relevant online consumer groups" (p. 62). When digital technologies and platforms are highly advanced, online communities become more open, and the fluidity of digital content crashes the boundaries of different communities through user interactions such as post sharing (Costello et al. 2017). The concept of communities becomes context-based and is defined depending on research purposes.

Netnography enables researchers to observe actor behaviours, such as word-of-mouth buzz on social media, or user discussion on online forums in a natural setting, and collect different types of user discourse such as texts, videos, photos or social interactions (Kozinets 2010). Netnography has been widely applied to examine diverse online marketing contexts. For example, Chua and Banerjee (2013) adopted netnography to investigate the social media marketing campaign of "My Starbucks Idea" and collected user posts from Twitter, Facebook and Foursquare, as well as from discussion boards. Compared to the methods such as interviews or focus groups, the data collected via netnography is less time-consuming and more representative of the moment of truth.

This thesis uses netnography as a pilot study to observe user interactions on social media and obtain a suitable research sample following Kozinets's (2010) six stages of netnography:

I. *Research Planning*: The validation of the DDSS framework was conducted using social media dialogue data. The service of social media customer care was chosen for investigating the issues related to complaint handling and corporate social innovation (CSI) embedded in social media dialogues.

II. *Entrée*: Twitter was chosen as the research platform. Twitter has been widely used by companies, and more than 70% of Fortune 500 companies operate an active Twitter account (Ratliff and Kunz 2014). Offering a microblogging service, Twitter allows only 140 characters in each post (known as a "tweet"), and thus Twitter dialogues are more similar to real-life dialogues between a customer and a service agent.

III. *Data Collection*: The data collection was conducted based on the Twitter platforms of four UK companies – a grocery retailer, a telecoms company, a bank, and a public transport service provider – during a six-month period. The researcher used open source software, Knime, to download both the companies' Twitter tweets and the customers' tweets sent to the companies. By matching the Twitter user ID mentioned in each tweet, dialogue streams could be identified.

IV. *Data Analysis*: The dialogue datasets were examined based on five criteria suggested by Kozinets (2010): research topic relevancy, sufficient number of postings, discrete message posters, rich content, and frequent between-member interactions. After evaluating the four datasets, the public transport company's dataset was excluded considering its low actor engagement and weak relevancy to the chosen topic – complaint handling and CSI.

V. *Ethical Standards*: Some ethical concerns are disclosed in this thesis regarding data collection. Firstly, whether user posts in cyberspace are private or public information has been an ongoing debate in prior research (Kozinets 2010). Often, social media users are unaware that they are being studied. Though researchers suggest that netnography should obtain online users' consent, it is unlikely to be executed in the big data research context where the number of identifiable users is vast. Secondly, the research topics regarding complaint handling and CSI may run potential risks of harming the companies' reputations, as the results tend to highlight poor business practices. Considering the potential ethical issues, this thesis anonymised the studied companies and their customers (or stakeholders) and removed sensitive information (e.g. user ID, personal information). After evaluating the content of the datasets, the bank dataset is considered unsuitable for further analysis since it contains sensitive personal details, especially bank account information.

VI. *Research Representation (Member Checks)*: Examining research representation is the process of presenting research findings back to research participants or relevant stakeholders to get feedback on the researchers' interpretations of data. Since netnography was used as a pilot study for collecting dialogue data and observing user interaction patterns, user communications, and dialogue structures, the representation assessment was conducted in the three applications during data analysis. The main approaches used were to engage relevant domain researchers in the focus groups and expert workshops to provide suggestions on the dialogue data analysis.

Finally, two Twitter dialogue datasets – the grocery retailer and the telecoms company – were used in this thesis for the validation of the DDSS framework. The details of the research datasets are provided in Table 2.5.

Table 2.5 Research Dataset

	The Telecoms Company	**The Grocery Retailer**
Total Number of Tweets	310,706	70,800
Number of Company Tweets	108,142	14,829
Number of Customer Tweets	202,564	55,971
Number of Dialogue	36,954	7,201

2.4.3.2 Three Applications of Social Media Dialogue Analysis

This thesis is presented in an alternative format, including three journal manuscripts. These papers describe three applications to evaluate the DDSS framework (design artefact) and demonstrate the feasibility of the proposed framework in addressing field problems. The proposed framework builds linkages between service systems, data environment, and decision-making and justifies the mechanisms of value co-creation occurring within them. The applications follow the DDSS framework to examine social media dialogue data in two different research contexts, and the findings from the applications are then fed back to the framework to improve the design artefact.

The first and second papers were carried out to examine service recovery via social media customer care (webcare). Complaint handling is one of the most important tasks of social media webcare to recover customer satisfaction after service failures (Grönroos 1988). Despite being a widely adopted practice, service recovery via social media is still in its infancy (van Noort et al. 2014), and the failure of service recovery tends to create negative social influence amongst complainers and other social media users (Schaefers and Schamari 2016). The first paper makes contributions to service recovery theory regarding dynamic value co-creation and offers insights into service recovery management. The second paper aims to provide a theorised dialogue model as well as an analytical method to improve the complaint dialogue mining. The results from the second paper enhance service recovery performance regarding company process recovery and customer recovery. Finally, the third paper was conducted to investigate social media dialogues in the context of corporate social innovation (CSI). Social media has become a "public arena of citizenship", where people talk about and share details of organisations' ethical practices. This paper proposes a data-driven approach to support CSI and offers operational guidance to extract and internalise stakeholder knowledge embedded in the dialogue data.

The full papers of the three applications are provided in **Chapters 3**, **4** and **5**.

Chapter 3 Understanding Service Recovery in a Dynamic Social Media Context Using Text Mining

ABSTRACT

Service recovery via social media creates a dynamic process of value co-creation as multi-actor engagements, customer co-recovery and interaction quality tend to affect recovery outcomes. This paper proposes, and empirically tests, a dynamic service-recovery framework based on social media dialogical interactions. Approximately 17,000 dialogues collected from a telecoms company's Twitter platform were analysed through a two-stage study. In Phase 1, text mining transformed dialogues into a structured format and identified customer pre-/post-recovery emotions, service failures, service recovery activities, and dynamic factors of recovery. In Phase 2, logistic regression was applied to model the factors impactful to service recovery. Empirical results indicate that service-recovery activities such as showing empathy and follow-up have important impacts on customer post-recovery emotions. In terms of the dynamic factors, a high level of competitor engagement and customer co-recovery leads to negative service-recovery experiences. Although the main effect of other users' engagement is not statistically proven, it is a key moderator of service recovery activities and outcomes. This study makes important contributions to service research by exploring diverse value co-creation interactions during service recovery. Importantly, this research provides crucial managerial implications by introducing a dialogue mining method, which enables companies to efficiently evaluate complaint-handling interactions and outcomes in the social media context.

Keywords: Service Recovery, Customer Complaint, Customer Experience, Social Media, Text Mining

3.1 Introduction

Social media has become an important channel of customer care. Increasingly, companies operate social media webcare to aggregate customer complaints and tackle service problems (Willemsen et al. 2013). The complaint handling actions are also known as service recovery, which is important to customer post-recovery satisfaction (Magnini et al. 2007), repurchase intentions (Grewal et al. 2008) and inclination to spread positive word-of-mouth (Davidow 2003; Maxham 2001). Unlike traditional channels, social media customer care operates in a dynamic service environment whereby all users are virtually present (Kaplan and Haenlein 2010) and can be passive or active participants in public dialogue streams (Baird and Parasnis 2011; Novani and Kijima 2012). Prior research has examined the effect of the presence of other social media users during service recovery (Schaefers and Schamari 2016). However, this has been done in a simulated research setting using questionnaires and taking user perceptions into consideration rather than as a field study using real-world data.

The dynamic service environment is also caused by customers' creation and co-creation of their unique service experience on social media (Baird and Parasnis 2011; Novani and Kijima 2012). Customers could choose if and when to engage and participate in value creation activities (Heinonen et al. 2010). This implies that, instead of involving customers in predefined service processes, companies should see service recovery as a interactive customer experience. Such interactions incorporate companies' recovery offerings and customer emotional responses. Customer emotions are a focal construct to consider in relation to whether they result in a positive or negative experience (Lemon and Verhoef 2016; Verhoef et al. 2009). Based on Twitter webcare, Fan and Niu (2016) discovered that the factors such as service agents' responses, speed of recovery and failure severity lead to a change in complainer emotions during service recovery. However, Fan and Niu's (2016) work only focused on the final status of customer emotions as the indicator of customer satisfaction, neglecting the process of emotional change, or more specifically, the quality of service-recovery interactions.

Moreover, previous research particularly investigated the co-created service recovery regarding customer participation in rectifying service problems, also defined as customer co-recovery (Dong et al. 2008; Roggeveen et al. 2012; Xu et al. 2014). The relevant research has examined how customer co-recovery differs from employee-initiated recovery in relation to customers' perceived justice and post-recovery satisfaction (Xu et al. 2014). However, prior studies were based on a simulated research context and did not consider the dynamic aspects of customer co-recovery, as these manifest themselves in a social media context, and where co-destruction outcomes may occur as the result of a lack of resources and knowledge (Xu et al. 2014).

In this paper, we analyse the dialogues of service recovery on a Twitter webcare platform. Dialogues are viewed as a value-enhancing interactive process of reasoning together (Grönroos 2004). Dialogue is an important component of co-creation as it requires the interactive parties' deep engagement, and the ability and willingness to act on one another (Prahalad and Ramaswamy 2004). Following Grönroos and Voimas' (2013) value-creation spheres, value co-creation takes place between a company and its customers and amongst customers. A conceptual framework is constructed to evaluate dynamic value co-creation during service recovery, with three dynamic components of multi-actor engagement, customer co-recovery and interaction quality. Previous studies have examined, to a limited extent, social influence (Schaefers and Schamari 2016) and customer co-recovery (Dong et al. 2008) during complaint handling. Our work makes important contributions to service research by building a linkage between service recovery literature and value co-creation. This study explores multiple value-co-creation interactions during service recovery and uncovers how the dynamic components affect companies' service-recovery effort.

Notably, social media data is regarded as one of the most important components of big data, relating to its characteristics of high volume, high variety, high velocity and high veracity (Kunz et al. 2017). Analysing a large number of service recovery dialogues on social media is a challenge for existing research. For example, Fan and Niu (2016) observed multiple Twitter platforms of airlines service providers in a five-month period but only investigated 347 pieces of dialogue. To address this issue, this research develops a novel data-analytical approach by proposing the ontology-based text mining to aiding the analysis of service-recovery dialogues. Ontologies are a semantic modelling tool for constructing domain knowledge that can be shared and re-used between people, or between people and machines (Lee et al. 2015; Mikroyannidis 2007). The proposed approach provides practical value, enabling researchers to perform semantic analysis, such as text mining, on large amounts of social media data, and to extract information effectively.

3.2 Conceptual Framework

Social media stimulates a high degree of user interaction, making it a potential game-changer in terms of service recovery. The fundamental changes are found in the ease of contact, volume, speed and nature of interactions (Gallaugher and Ransbotham 2010). Larivière et al. (2013) advocated that dialogical interactions among social media users promote information value (e.g. online reviews, word-of-mouth), identity value (e.g. expressing personality and status), social value (e.g. gaining social approval) and monetary value (e.g. choosing between competing offers). Despite the opportunities arising from nurturing a social media platform for customer engagement (Gallaugher and Ransbotham 2010), the dynamic social media environment hampers the ability of companies to effectively manage the social dialogues. In this paper, we conceptualise the dynamic dialogical interactions of service recovery in three aspects: multi-actor engagement, interaction quality and customer co-created recovery.

3.2.1 Multi-Actor Engagement in Customer Complaints

The growth of social media platforms enables customers to share their negative service experiences more easily (Grégoire et al. 2014; Pfeffer et al. 2014). Differing from traditional channels, social media transforms the one-to-one service process into a collective customer experience (Baird and Parasnis 2011; Novani and Kijima 2012). Hence, customer complaints on social media inevitably become a co-created experience among multiple actors.

Grégoire et al. (2014) outlined six types of social media complaining behaviour: directness (directly contacting the company); badmouthing (negative word-of-mouth without interacting with the company); boasting (positive publicity about exceptional service recovery); tattling (complaining to a third party); spite (spreading negative publicity to get revenge); and feeding the vultures (enabling competitors to benefit from complaints). Einwiller and Steilen (2014) indicated that complainers have different motivations. Alongside personal goals, such as anxiety reduction and solution seeking, collective goals such as altruism are a critical driver of customer complaining behaviour. Social media enables complainers to achieve both personal and collective goals; that is, reaching service providers with a virtual presence (Grégoire et al. 2014) and turning complaints into a wider public action through which complainers can potentially change one another's perceptions and actual behaviour (Tripp and Grégoire 2011; Libai et al. 2010).

Importantly, the multi-actor interactions may amplify the degree of severity of the original complaint through information sharing (e.g. online firestorm, as defined in Pfeffer et al. 2014). Moreover, competitors can also participate and "gate-crash" through social media interactions. Larivière et al. (2013) stressed that social media provides a networking environment whereby customers can instantly acquire competitors' information and offerings. Often, social media transforms customer complaints into a market battle between companies. Instances have been reported in the press; for example, when the mobile phone brand Nokia leveraged a Twitter customer complaint for Samsung by offering its brand new mobile phone to the complainer. Despite some prior research anchoring in the area of competitors taking advantage of social media complaints (Grégoire et al. 2014), little attention has been paid to the issue of competitors' real-time engagement in service recovery.

3.2.2 Interaction Quality of Service-Recovery

Interaction quality has been highlighted in early research, referring to customers' perception of the process and manner in which services are delivered by service encounters (Grönroos 1984, 1988). Interaction quality has a direct impact on customer satisfaction and customer inclination to maintain the relationship with service providers and it is often evaluated by employee-related attributes such as service agents' politeness, friendliness, sensitivity, and empathy (Choi and Kim 2013; Ekinci and Dawes 2009).

To understand service-recovery interactions, justice theory is the most widely adopted theoretical framework (Buttle and Burton 2002). In general, justice theory postulates that the customer's feeling of justice can be recovered depending on the ongoing evaluation of their losses from the service failure and the gains from the recovery provisions (Tax et al. 1998). The concept of perceived justice is assessed by three elements, namely, distributive justice, procedural justice and interactional justice (Smith et al. 1999). Distributive justice focuses on the outcomes that a complainer receives, such as compensation, refunds and problem correction, whereas procedural justice relates to a company's complaint-handling process (Smith et al. 1999). Interactional justice is further divided into interpersonal justice and informational justice (Lee and Park 2010). Interpersonal justice is relevant to customers' judgement of frontline employee manners, attitudes and competence, whereas informational justice refers to the quality of the explanation that customers receive about decision-making (Ambrose et al. 2007). Notably, the justice dimensions are not independent and there are compounded effects between different types of justice (Goodwin and Ross 1992).

Although justice theory serves as a useful framework allowing companies to define and evaluate service recovery (Buttle and Burton 2002), it does not clarify customers' role throughout the service interactive processes. Service recovery via social media should not be a pre-designed service activity but a co-created experience (Cheung and To 2016). Customer–company interaction forms a unique customer experience, during which a company's service offering will trigger a customer outcome, be it emotional or behavioural (Verhoef et al. 2009). Heinonen et al. (2010) stressed that such a customer experience is dynamic and uncontrollable for companies, as customers can choose to enter, participate and exit at any point of value co-creation.

Prior research pointed out that customers' actual behaviour is commonly emotion-driven, and thus, companies should deal with customer emotions first and service problems second (Chebat and Slusarczyk 2005). Recent business practice has also emphasised the importance of coaching frontline employees to resolve the emotional side of customer interaction through avoiding the use of words that may trigger negative feelings, such as "can't", "won't" and "don't" (Dixon et al. 2010). Moreover, previous research argues that the driver of complaining behaviour might not be the dissatisfaction per se, but that it might also be triggered by the antecedent negative emotion arising from the disappointing service experiences (Tronvoll 2011). Recovering customer perceived justice can also create positive emotions (Ozgen and Kurt 2012). In addition, it has been found that customer emotional responses mediate the effect of customer perceived justice on post-recovery satisfaction in several service-recovery contexts (Chebat and Slusarczyk 2005; del Río-Lanza et al. 2009; Ozgen and Kurt 2012).

Collins (1981), from a sociological perspective, viewed emotions as a type of resource that is gained from successful social interactions (positive emotions) and decreases through undesirable interactions (negative emotions). Individuals go through a chain of interactions, matching conversational and emotional resources obtained from past encounters until their emotion stabilises or declines (Collins 1981). Emotional responses are also a key ingredient of the customer service experience (Verhoef et al. 2009). Customers assess the service performance depending on numerous service clues and, in turn, the clues influence both the rational and emotional judgement of customers regarding service quality (Berry et al. 2006). Considering interaction quality as a key to triggering emotional change, it is important to identify the service activities that positively affect customer emotions.

3.2.3 Customer Co-Recovery

In addition to company-initiated service recovery, Meuter and Bitner (1998) identified two other types of service recovery based on the degree of customer participation, namely, joint recovery and customer recovery. Customer recovery emphasises the service recovery performed by customers themselves, other customers or a third party requested by the complainer (Meuter and Bitner 1998). The customer-recovery actions are known as customer-to-customer (C2C) recovery (Nicholls 2010). C2C recovery is commonly found in virtual communities where users seek support or troubleshooting guidance from others.

In terms of joint recovery, customers are viewed as "partial employees" who contribute time and other resources to undertake a part of the service-recovery functions under employees' guidance (Dong et al. 2008; Meuter and Bitner 1998). Dong et al. (2008) extended the concept of joint recovery and coined the term co-recovery to capture customer value co-creation in the complaint-handling process. Xu et al. (2014) explained that customer co-created recovery is a solution generated through interactions between the company and its customers. During the interactions, customers, frontline employees and managers input their knowledge, skills, time and other resources to co-create value (Edvardsson et al. 2011).

The co-creation of service recovery is found to positively impact the service-recovery outcomes. Customers possessing higher degrees of service-recovery participation are more likely to perceive justice and regain greater satisfaction (Dong et al. 2008; Roggeveen et al. 2012). Successful customer co-recovery can also create a higher tendency towards customer repurchasing in the future (Xu et al. 2014). Thus, companies need to satisfy customers' needs for co-recovery (Roggeveen et al. 2012). It has been suggested that more research is needed to understand co-recovery by exploring which service contexts and mechanisms alter the corresponding impact of co-recovery, as well as actor participation (Dong et al. 2008). Especially in the context of social media, the virtual interaction requires a higher level of customer engagement in co-recovery.

3.3 Research Questions and Design

To evaluate the three dynamic components of service recovery – engagement of multiple actors, customer co-recovery and interaction quality, we need to examine the types of service-recovery activity that are delivered through social media in relation to customer emotional responses and, specifically, the ones creating positive customer emotions. On the basis of these, we specify the following five research questions:

RQ1 What types of service-recovery activities relating to the different justice perspectives can be delivered on social media?

RQ2 What types of service-recovery activities can create positive customer emotions?

RQ3 How does the engagement of multiple actors on social media affect service-recovery experiences?

RQ4 How does customer co-recovery affect service-recovery experiences?

RQ5 How can the service-recovery experience be examined from an interaction perspective?

These research questions were addressed by a two-stage study, as shown in Figure 3.1. In Phase 1, RQ1 was addressed by employing an ontology-based semantic method to analyse real-world dialogical interactions, extracting information concerning service failures, service-recovery activities and changes in customer emotions throughout recovery interactions. Phase 2 was conducted to address RQ2, RQ3, RQ4 and RQ5. We analysed the direct and moderating effect of the dynamic constructs on influencing service-recovery experience. The proposed text analytics method improves complaint management via social media, and research findings contribute to both service-recovery theory and practice by providing guidance for managing the service-recovery experience more effectively.

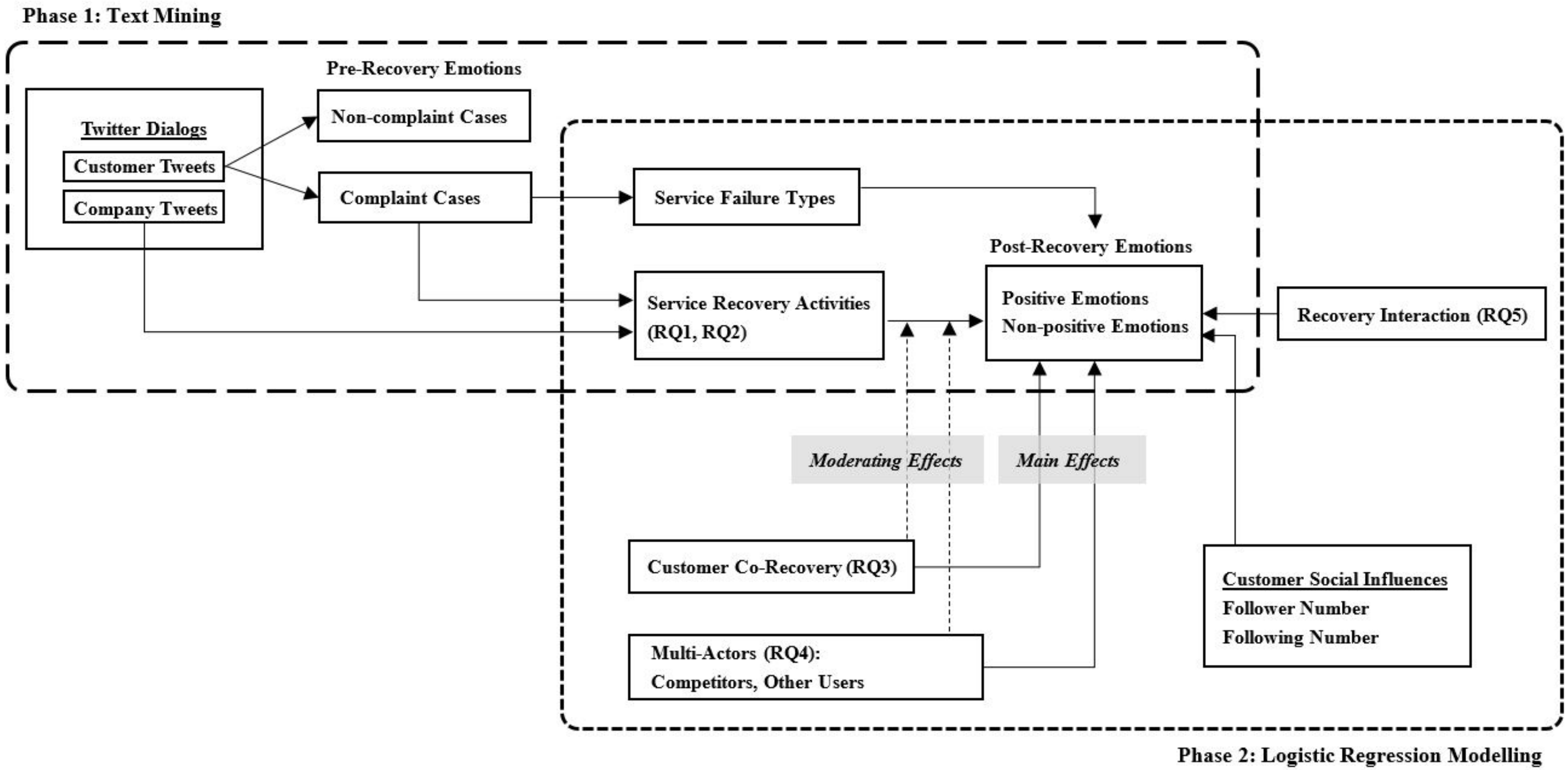

Figure 3.1 Research Design

3.4 Phase 1 Text Mining

3.4.1 Analysis Method

We collected real-world dialogues from a UK telecommunications company's Twitter webcare over a six-month period. Solis and Kutcher (2011) noted that Twitter has become a critical social media platform, with more than half of the Fortune 100 companies using Twitter for customer service, online community management and information diffusion. Possessing companies' high level of social presence, Twitter is considered a suitable research platform. In analysing Twitter dialogues, we developed a four-stage text-mining process, including dialogue identification, sentiment analysis, ontology building and information extraction.

3.4.1.1 Stage One: Dialogue Identification

Twitter serves as a microblogging service platform on which only 140 characters are allowed in each post, known as a "tweet". A Twitter user can interact with others through three types of communication – "mentioning" another user in a post, "replying" to other users' posts and "retweeting" others' posts. In the data-collection period we gathered 310,706 tweets from the company's Twitter webcare, including 202,564 customer tweets to the company from 72,084 unique users and 108,142 company reply tweets. The telecoms company's Twitter platform satisfies Kozinets' (2002) five criteria for online data collection, namely, research topic relevancy, high traffic of postings, large volume of distinct message posters, rich data and a high degree of between-member interactions. Thus, it is regarded as a suitable data set for further analysis.

The Twitter data set contains rich information, such as *"User id"*, *"Tweeting time"*, *"Tweet content"* and users' social profiles (e.g. location and registering time).

In order to capture company–customer interactions, we examined tweet dialogues containing a single customer tweet to the company, and the company's replies, by matching user IDs. In this phase, a total of 36,954 dialogues or "mini-cases" were identified. Customer tweets and also company tweets without interactions were excluded.

3.4.1.2 Stage Two: Sentiment Analysis

Sentiment analysis examines the polarity of a document, uncovering people's feelings about a topic as positive, negative or neutral (Xu et al. 2011). Opinion-mining research sees negative sentiment hosting customer negative emotions as a clue to customer complaints (Liau and Tan 2014). We employed sentiment analysis to aid complaint extraction from the twitter dialogues. Since this study aims to understand changes in customer emotions in the service-recovery experience, we need to examine all types of sentiment. Schoefer and Ennew (2005) suggested that customers' positive emotions are directly related to customer satisfaction after superior service recovery. Often, customers tend to spread positive WOM messages after a successful service recovery (Maxham 2001). Therefore, positive sentiment is viewed as an ideal indicator of a positive service-recovery experience.

There are numerous commonly shared sentiment lexicons available for analysing customer emotions, such as SentiWordNet (available from http://sentiwordnet.isti.cnr.it) and MPQA Opinion Corpus (available from http://www.cs.pitt.edu/mpqa/). However, none of these is tailored to Twitter communications – in other words, they do not consider emoticons – and thus it was decided to create a specific sentiment lexicon to extract customer emotions. To do this, a random number of 1,000 Twitter dialogues was used for manual annotation by two domain experts in service research area, following Smith and Bolton's (2002) five negative emotions (e.g. "pissed off" coded in to *Angry*), and specific aspects of Gregg and Scott's (2008) complaint typology (e.g. "on hold for 20 minutes" coded into *Wait Time*). Moreover, we also coded positive words as positive sentiment (e.g. "excellent customer service" coded into *Positive Service*).

3.4.1.3 Stage Three: Ontology Building

In addressing RQ1, a domain ontology is useful to examining customer complaints and service-recovery activities embedded in social media dialogues. Ontologies provide a shared conceptualisation that demonstrates the world's phenomena by identifying concepts and inter-relationships used for semantic analysis (Cao et al. 2011; Gruber 1993). Applying ontologies to understand text structures can help researchers to build up a shared model that is available for communication and re-use between humans, software systems, or between humans and software systems (Mikroyannidis and Theodoulidis 2010). In marketing research, ontologies have been applied to demonstrate the linkage between different types of antecedent and customer response in the complaining process (e.g. Crié 2003), yet the use of ontologies to explain service recovery remains under-explored (e.g. Lee et al. 2015). We employed a three-layer ontology structure proposed by Missikoff et al. (2002), which includes top-layer, upper-layer and domain-specific ontology. The degree of abstraction increases from the bottom layer (domain-specific ontology) to the upper layer (top-layer ontology). Such a multi-layer structure contains ontologies for specific purposes at each layer and, in addition, should include the intra-layer and inter-layer relationships between concepts (Mikroyannidis and Theodoulidis 2010).

The domain experts conducted manual annotation on the 1,000 Twitter dialogues that had been used in customer-emotion coding. The coding procedure was designed to capture key terms and phrases used in both customer and company tweets. We used the ARC (activity, resource and context) framework developed by Ordenes et al. (2014) for the top-layer ontology. The ARC framework was designed to analyse unstructured customer feedback based on service logic. Similar to the methodology of Ordenes et al. (2014), we created the upper-layer ontology by identifying the actors (e.g. customers, the company, third-party organisations) who possess resources and participate in activities. In the bottom-layer ontology, we detailed the resources and activities mentioned in the tweets. An example of Twitter dialogue coding is provided in Table 3.1. As shown, in the customer tweets we identified terms relating to company activities or resources (e.g. telecoms service, customer support service) and customer activities or resources (e.g. time). Notably, customer emotions, coded in Phase 2, are viewed as a customer resource that can be gained or lost during service interactions.

In the company reply tweets we identified terms relating to recovery activities following justice theory (Smith et al. 1999; Tax et al. 1998). We captured concepts of service-recovery activities, such as showing empathy (interpersonal justice) and solution guidance (distributive justice). Particularly, in the conditions whereby service-recovery activities require customer-resource engagement, such as “drop us your postcode”, we coded these into customer co-recovery activity (see Table 3.1). In total, we classified eight categories of customer resource, eleven categories of customer activity, thirteen categories of company resource, thirteen categories of company activity and three contexts related to competitors and other companies.

Ontology development is an iterative process (Missikoff et al. 2002). In order to evaluate the ontologies, a focus group was used to engage domain experts to assess the primary ontologies and examine whether the concepts were classified appropriately (Haghighi et al. 2013). Eight coders with a marketing background participated in the focus group as domain experts. They were separated into two groups: a complaint-coding group and a recovery-coding group. Each group worked on the customer tweets and the company reply tweets, respectively. Within each group, each pair of coders was given 200 tweets and was asked to code these using the ontologies. Two judges compared the agreement between coders and examined the conflicting annotations. The concepts with the disagreement were re-evaluated by the judges and the ontology was revised accordingly. In total, 3,588 terms were captured in the manual annotation process. The structure of ontologies are provided in Appendix A.

Table 3.1 Example Manual Annotation of Twitter Dialogues

	Terms	Domain-Specific Ontology	Upper-Layer Ontology	Sentiment
Customer Tweet	NO 3G			Service Failures
@Telecom Company I've had NO 3G or 4G for the past 2 days!! Pissed off! #PLEASEHELP!	Pissed off!	Emotions	Customer Resource	Negative Emotions: Angry
	3G or 4G	Telecom Service	Company Resource	
	past 2 days	Time	Customer Resource	
	PLEASEHELP!	Customer Support Service	Company Activity	
Company Reply Tweet	oh no :-(	Interpersonal Justice: Showing Empathy	Recovery Activity	
@Customer Hi Ben, oh no :-(Sorry about that. Drop us your postcode and we'll take a look. Or chat with us here: (weblink)	Sorry about that	Interpersonal Justice: Apology	Recovery Activity	
	Drop us your postcode	Customer Co-Recovery	Recovery Activity	
	we'll take a look	Distributive Justice: Problem Handling	Recovery Activity	
	chat with us here:	Procedural Justice: Channel Directing	Recovery Activity	

3.4.1.4 Stage Four: Information Extraction

Text mining reduces the complexity of a collection of documents by using a set of concepts to represent these documents, thereby transforming unstructured data into a structured format for further analysis (Ordenes et al. 2014). By applying the sentiment lexicon and ontologies developed in the previous two stages, we built up text analysers (called libraries) using IBM SPSS Modeler and extracted the information about customer complaints and the company's service recovery from 35,954 Twitter dialogues.

The information extraction stage comprised three tasks (see Figure 3.1): first, customer complaints were extracted by identifying negative sentiments embedded in customer tweets. We specified the corresponding categories of service failure by linking customer negative sentiments to certain types of resource or activity of the company, customers and other actors. For instance, a tweet classified as a Telecoms Issue might include the company's telecoms resource "3G" or "4G" and the customer's negative emotion "pissed off" (see Table 3.1). In this way, we conceptualised three types of service failure, namely, service value failure (company-initiated failure), co-creation failure (customer-initiated failure) and uncontrollable failure (third-party-initiated failure).

Second, we identified successful service-recovery cases by examining the sentiment changes from negative (customers' first tweet) to positive (customers' final tweet). Often, positive sentiment contains customer sarcasm, which is a well-documented issue in opinion-mining research (Davidov et al. 2010). The suggested method to detect sarcasm is to recognise the contrasting contexts, in which a positive sentiment document simultaneously contains negative words (Riloff et al. 2013). Following this approach, we removed cases with contrasting sentiment. Finally, in the identified customer complaint tweets, we examined the corresponding company replies and extracted service recovery activities and customer co-recovery activities.

3.4.2 Results

Using text-mining method, we identified 17,125 complaint cases, which include 14,128 cases referring to specific service failures and 2,997 cases containing only negative emotions without any resources or activities mentioned (see Table 3.2). In addition, a total of 2,024 complaint cases were found to contain a positive service-recovery experience; this represents a success rate of 11.82 per cent in the sample.

The three main types of service failure, namely, service value failure, co-creation failure and uncontrollable failure, are shown in Table 3.3 Service value failures are the services that a company promises, yet fails, to deliver. There are 11 sub-categories of service value failure, and 52.92 per cent of the cases belong to *Telecoms Issue*. Co-creation failure refers to failures caused by customers themselves (or their resources) when using the service (Huang 2008), such as “run out of data”, “lost my phone”. There were 291 (1.7%) co-creation failure cases identified. Finally, uncontrollable failures relate to service problems caused by a third party (Huang 2008), such as *Fraud* (0.43%) or *Other Companies’ Failure* (0.17%).

Table 3.2 Summary Statistics

	Doc. Count	Percentage
Pre-Recovery		
Service Failures	14,128	82.50%
Negative Emotions Only	2,997	17.50%
Post-Recovery		
Positive Recovery Experience	2,024	11.82%
Non-Positive Recovery Experience	15,101	88.18%
All Cases	17,125	100%

Table 3.3 Breakdown of Service Failures

	Doc. Count	Percentage	Accuracy	Precision	Recall
Service Value Failure					
Telecoms Issue	9,063	52.92%	96%	92%	100%
Incompetence	2,762	16.13%	89%	82%	95%
Contacting Company	2,308	13.48%	95%	90%	100%
Waiting Time	1,788	10.44%	88%	76%	100%
System Malfunction	1,183	6.91%	83%	70%	94%
Price and Payment	1,062	6.20%	89%	78%	100%
Mistreatment	905	5.28%	89%	80%	97%
Mistake	706	4.12%	83%	74%	97%
Delivery and Not Receiving	443	2.59%	88%	76%	100%
Product and Stock	390	2.28%	83%	90%	97%
Promotion and CRM	398	2.32%	91%	82%	100%
Co-creation Failure					
Customer-Initiated Failure	291	1.70%	86%	72%	100%
Other Uncontrollable Failure					
Fraud	74	0.43%	81%	62%	100%
Other Companies' Failures	29	0.17%	100%	100%	100%

Table 3.4 shows the text-mining results of service-recovery activities. A total of 11 types of service-recovery activity was extracted. The most frequent activity was *Showing Empathy* (71.15%). The company showed empathy to complainers by replying with written response and emoticons, such as “oh no :(” or “that's not good”,

to comfort complainers' negative emotions. The second highest-frequency activity was *Solution Guidance* (40.79%); that is, instructing complainers in solving service problems by mentioning terms such as "reboot your phone" to help with phone setting management, or "Have you tried…?" to suggest other actions. Unsurprisingly, the least frequent activity was *Refund & Replacement* (0.89%).

In addition to company-initiated service recovery, we also extracted the service-recovery cases with customer participation. There are 7,546 *Customer Co-Recovery* cases (44.06 %) being found in the Twitter dialogues. Notably, *Customer Co-Recovery* often co-occurred with other company-initiated activities, especially *Solution Guidance* in which customers directly conducted trouble-shooting under the company's instruction and *Information/Explanation* in which customers were requested for providing their knowledge, such as "do you know other users having this issue?" to help the company understand problematic situations. A further examination on how *Customer Co-Recovery* affect company-initiated service recovery will be discussed in Phase 2.

Overall, the automatic information extraction performed well. To evaluate the performance of text mining, we applied the widely used performance measures, namely, precision, recall and F-measure (Manning et al. 2008). For service failures, we obtained accuracy of between 100 and 81 per cent, and 89.3 per cent accuracy in successful recovery cases. In terms of recovery activity extraction, accuracy was between 100 and 87 per cent. However, around 5% of company reply tweets had no concepts extracted by our text analysers. The number of unrecognised cases are considered reasonable, as the real-life dialogues contain short responses, such as "Yes", "Hmm…" and noise, which cannot be assigned to a service-recovery category.

Table 3.4 Summary of Service-Recovery Activity Extraction

Recovery Activities	Doc. Count	Percentage	Accuracy	Precision	Recall
Company Recovery					
Showing Empathy	12,184	71.15%	100%	100%	98%
Solution Guidance	6,985	40.79%	87%	80%	93%
Follow-Up	4,579	26.74%	88%	76%	100%
Problem Handling	4,499	26.27%	90%	82%	97%
Information/Explanation	4,303	25.13%	88%	84%	91%
Apology	4,150	24.23%	100%	100%	100%
Active Help	3,667	21.41%	100%	100%	100%
Channel Directing	2,935	17.14%	95%	94%	95%
Problem Acknowledgement	1,240	7.24%	87%	76%	97%
Guarantee	1,112	6.49%	94%	92%	95%
Refund & Replacement	153	0.89%	90%	80%	100%
Customer Co-Recovery	7,546	44.06%	94%	96%	92%
Uncategorised concepts	863	5.03%	-	-	-

3.5 Phase 2 Logistic Regression

Phase 2 builds on Phase 1 and examines the effects of important variables on the service-recovery experience (see Figure 3.1). Through Phase 1, the unstructured Twitter dialogues were transformed into a structured data set using text mining. In this phase, we conducted two experiments and modelled the variables – service failures, service-recovery activities, and customer emotions change – to address the research questions. Experiment 1 examined RQ2, RQ3 and RQ4, whereas experiment 2 examined RQ5.

3.5.1 Experiment 1

In Experiment 1 we first analysed the main effects of service failure types and service-recovery activities on service-recovery experience. Subsequently, the two dynamic constructs (multi-actor interactions and customer co-recovery) were then examined to understand their main effects on the service-recovery experience and moderating effects on the relationships between company-initiated recovery activities and service-recovery experience. Logistic regression was employed for data analysis. Logistic regression has been widely used to predict customer behaviour such as customer churn (Neslin et al. 2006) and customer referral behaviour (Stein and Ramaseshan 2015). Different from linear regression, the logistic regression model aims to test a dichotomous outcome variable, investigating non-linear relationships (King 2008). The logistic regression model summarises the data in terms of contributing or not contributing to the outcome variable, based on a combination of values of the inputting predictors. The outcome variable can be interpreted as an odds ratios based on the conditions where predictors present or change (Hosmer et al. 2013).

Main Effects

To analyse the main effects of service failure types (*SF*), company recovery activities (*RA*), customer co-recovery (*CR*), and multi-actor engagement (*MAc and MAo*) on the service-recovery experience (*SRE*), we used the following models:

$$\mathrm{SRE}_i = \beta_1 + \beta_2\mathrm{SF}_i + \beta_3\mathrm{RA}_i + \beta_4\mathrm{Inter}_i + \beta_5\mathrm{Soc}_i + \varepsilon_i \tag{Eq. 1a}$$

$$\mathrm{SRE}_i = \beta_1 + \beta_2\mathrm{SF}_i + \beta_3\mathrm{RA}_i + \beta_4\mathrm{Inter}_i + \beta_5\mathrm{Soc}_i + \beta_6\mathrm{MAc}_i + \varepsilon_i \tag{Eq. 1b}$$

$$\mathrm{SRE}_i = \beta_1 + \beta_2\mathrm{SF}_i + \beta_3\mathrm{RA}_i + \beta_4\mathrm{Inter}_i + \beta_5\mathrm{Soc}_i + \beta_6\mathrm{MAo}_i + \varepsilon_i \tag{Eq. 1c}$$

$$\mathrm{SRE}_i = \beta_1 + \beta_2\mathrm{SF}_i + \beta_3\mathrm{RA}_i + \beta_4\mathrm{Inter}_i + \beta_5\mathrm{Soc}_i + \beta_6\mathrm{CR}_i + \varepsilon_i \tag{Eq. 1d}$$

In all the above equations (Eq. 1a–1d), SRE_i is defined as the emotions of complainer i after service recovery, which was identified in Phase 1 using text mining (see Table 3.2). Each service-recovery dialogue would only take one of two binary conditions: positive or non-positive (negative or neutral) recovery experience. It is worth noting that the positive cases account for only 11.82 per cent (2,024 cases) of this data set and the majority of Twitter service recovery is non-positive cases (15,101 cases, 88.18%). More specifically, the data set is imbalanced, which challenges the model fits (Dawes 2009). Furthermore, we included SF_i and RA_i, which represent 14 types of service failure (see Table 3.3), and 11 types of company-initiated recovery activity (see Table 3.4), respectively. Dawes (2009) suggested that covariates should be included in the logistic regression model to ensure respondent heterogeneity and avoid confounding impacts on the results. In the Twitter data set, we further created three variables. These are Inter_i (the frequency of dialogical interaction between a complainer and the company), Soc_i – *Twitter Follower Number* (the number of other Twitter users subscribing to a user's posts) and *Twitter Following Number* (the number of other users' Twitter feeds subscribed to by a Twitter user). The descriptive statistics of the predictor variables are provided in Appendix B.

Eq. 1a represents a base model that includes the covariates and predictor variables to identify the main effects between the criterion variable and other predictor variables. Then, the dynamic constructs were added to the base model. MAc_i and MAo_i are the two predictor variables modelling multi-actor participation in service recovery, and they represent *Competitor Participation* (MAc_i) in Eq. 1b – whether a Twitter post has competitors' engagement and *Other Users' Involvement* (MAo_i) in Eq. 1c – whether a Twitter post has other users' (non-competitors) involvement. Both MAc_i and MAo_i are numerical values showing the frequency of

nvolvement. Both MAc_i and MAo_i are numerical values showing the frequency of

MAc_i and MAo_i are numerical values showing the frequency of user presence in a Twitter dialogue. Finally, in Eq. 1d, CR_i is added to the base model; this variable represents *Customer Co-recovery*, which created in the text-mining phase (see Table 3.4).

We performed univariable logistic regression model suggested by Hosmer et al. (2013, p.91) on each predictor variables as a pre-test and as a feature selection method to identify important features from the 31 predictor variables. The results of the pre-test indicate that 16 variables demonstrate significant differences in the two groups of positive and non-positive recovery experience ($p < 0.05$, using the Wald statistic). They are four categories of service failure (*SF: Telecoms Issue, System Malfunction*, *Mistreatment* and *Incompetence)*; nine categories of the company's recovery activities (*RA: Solution Guidance*, *Problem Handling*, *Channel Directing*, *Information/Explanation*, *Follow-Up*, *Apology*, *Active Help*, *Showing Empathy,* and *Customer Co-Recovery*); *Competitor Participation* - *MAc*; *Interaction Frequency* (*Inter*); and *Following Number* (*Soc*). Although the variable Other Users' Involvement – MAo was not identified as an important feature in the pre-test ($\beta = 0.003$, $p = 0.77$), we still included it in the model considering that MAo may possess a moderating effect on service recovery outcome. We used these 17 variables as inputs to perform the logistic regression models and excluded the non-significant variables.

Moderating Effects

In addition to investigating the main effects, we also analysed the moderating effect of multi-actor engagement and customer co-recovery. We applied a moderated regression analysis to investigate whether the relationship between a predictor variable and a criterion variable is moderated by the third variable (Sharma et al.

1981; Stein and Ramaseshan 2015). The moderated variable contributes to both theory and practice by setting up boundary conditions for the relationships of interest (Aguinis 2004, p.4). To test the moderating effects, we used the following models:

$$\mathrm{SRE_i} = \beta_1 + \beta_2\mathrm{SF_i} + \beta_3\mathrm{RA_i} + \beta_4\mathrm{Inter_i} + \beta_5\mathrm{Soc_i} + \beta_6\mathrm{MAc_i} + \beta_7\mathrm{RA_i}\times \mathrm{MAc_i} + \varepsilon_i$$

Eq. 2a

$$\mathrm{SRE}_i = \beta_1 + \beta_2\mathrm{SF}_i + \beta_3\mathrm{RA}_i + \beta_4\mathrm{Inter}_i + \beta_5\mathrm{Soc}_i + \beta_6\mathrm{MAo}_i + \beta_7\mathrm{RA}_i\times \mathrm{MAo}_i + \varepsilon_i$$

Eq. 2b

$$\mathrm{SRE}_i = \beta_1 + \beta_2\mathrm{SF}_i + \beta_3\mathrm{RA}_i + \beta_4\mathrm{Inter}_i + \beta_5\mathrm{Soc}_i + \beta_6\mathrm{CR}_i + \beta_7\mathrm{RA}_i\times \mathrm{CR}_i + \varepsilon_i$$

Eq. 2c

In Eq. 2a and Eq. 2b, the multiplicative interaction terms are used to evaluate the moderating effects of multi-actor engagement, including *Competitor Participation* ($\mathrm{RA}_i\times \mathrm{MAc}_i$) and *Other Users' Involvement* ($\mathrm{RA}_i\times \mathrm{MAo}_i$), on the relationships between company-initiated recovery activities and service-recovery experience. Similarly, the term $\mathrm{RA}_i\times \mathrm{CR}_i$ in Eq. 2c describes the moderating effect of customer co-recovery on the main effect of company-initiated recovery activities.

Experiment 1 Results

An examination of the correlations among the predictor variables was conducted, showing the absence of multicollinearity in the predictor variables (see Appendix C). The covariate *Interaction Frequency (Inter)* was found to possess a moderate correlation to *RA* activities *Solution Guidance* and *Showing Empathy;* yet for most of the predictors, it remains weakly correlated. Thus, we ignored this when performing the logistic regression analysis.

Table 3.5 illustrates the four models pertaining to the main effects models Eq. 1a to Eq. 1d. Model 1a represents the base model without multi-actor interactions and customer co-recovery corresponding to Eq. 1a. As shown in Table 3.5, most predictor variables have a negative impact on *SRE* outcomes, including *SF* variables *Telecoms Issue* (β = -0.236, $p < 0.01$), *Mistreatment* (β = -0.494, $p < 0.01$), *Incompetence* (β = -0.224, $p < 0.01$) and *RA* variables *Solution Guidance* (β = -0.105, $p < 0.01$), *Apology* (β = -0.326, $p < 0.01$) and *Showing Empathy* (β = -0.414, $p < 0.01$). Positive impacts were found for *Interaction Frequency* (*Inter*) (β = 0.294, $p < 0.01$) and *SF* variables *System Malfunction* (β = 0.169, $p < 0.05$) and *Follow-Up* (β = 0.378, $p < 0.01$). Similar results can be seen in Models 1b, 1c and 1d. For example, *Showing Empathy* is significantly supported in the four models with similar coefficients (β = -0. 414 in model 1a, β = -0.435 in model 1b, β = -0.415 in model 1c, β = -0.414 in model 1d).

The resulting logistic regression models also show that the dynamic constructs *MAc* (Model 1b), *MAo* (Model 1c) and *CR* (Model 1d) have a negative influence on the service-recovery experience (*SRE*). More specifically, the dynamic variables decrease the odds ratio of successful service recovery. Notably, the empirical results only support the effects of *MAc* (β = -0.207, $p < 0.01$) and *CR* (β = -0.153, $p < 0.01$), while *MAo* is not statistically significant (β = -0.004, p = 0.778).

Table 3.6 shows the results of the moderating effects of *MAc* (see Eq. 2a), *MAo* (see Eq. 2b) and *CR* (see Eq. 2c). As shown in Model 2a, no findings were statistically significant for *MAc* in moderating the effects of the company's recovery activities (*RA*) on the service-recovery experience (*SRE*). In terms of *Mao*, its main effect was found not to be significant, but it had negative moderating effects on *RA*.

Model 2b shows the negative signs in the interaction term *MAo* × *RA* (*Problem Handling*) ($\beta = -0.083$, $p < 0.05$) and *MAo* × *RA* (*Follow-up*) ($\beta = -0.088$, $p < 0.05$). This implies that the more other users are involved in these two *RA* activities, the less the likelihood of a positive recovery experience (*SRE*).

For *CR* (see Model 2c), negative signs were found in the interaction terms *CR* × *RA*(*Channel Directing*) ($\beta = -0.103$, $p < 0.05$), *CR* × *RA*(*Apology*) ($\beta = -0.135$, $p < 0.01$), *CR* × *RA*(*Active Help*) ($\beta = -0.147$, $p < 0.01$), and *CR* × *RA*(*Follow-up*) ($\beta = -0.152$, $p < 0.01$). In contrast, *CR* had a positive moderating effect on the relationships between *RA* (*Showing Empathy*) and *SRE* ($\beta = 0.087$, $p < 0.01$). Notably, in Model 2c the interaction parameters and predictor variables both had effects ($p < 0.05$) on *SRE*. Dawes (2009) explained that in this condition *CR* should be viewed as a quasi-moderator.

All the tested models were found to be significant by χ^2 analysis ($p < 0.01$). Three fit statistics were used to assess logistic regression models. These are –2 log likelihood ratio, Cox & Snell R Square and Nagelkerke R square. Overall, the inserted variables of *MAc*, *Mao* and *CR* barely created improvement on the model fit. The Nagelkerke R square was around 0.07 for all models in Table 3.5 and Table 3.6. This is because the data set was imbalanced and the cases of positive service-recovery experience only accounted for approximately 11 per cent. In other words, the naïve logistic model with only the intercept can accurately classify approximately 88 per cent of cases by simply attributing each case to a non-positive service-recovery experience (Dawes 2009). This is a well-acknowledged issue in logistic regression modelling.

Table 3.5 Main Effects of Competitors, Other Twitter Users and Customer Co-Recovery

	Model 1a			Model 1b			Model 1c			Model 1d		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Constant	-2.175**	.046	.114	-2.131**	.046	.119	-2.173**	.046	.114	-2.195**	.046	.111
Inter	.294**	.019	1.341	.306**	.019	1.359	.294**	.019	1.342	.333**	.021	1.395
Soc(Following Number)	.000	.000	1.000	.000*	.000	1.000	.000	.000	1.000	.000	.000	1.000
SF(Telecoms Issue)	-.236**	.029	.790	-.227**	.030	.797	-.235**	.030	.790	-.231**	.029	.794
SF(System Malfunction)	.169*	.074	1.184	.151*	.074	1.163	.169*	.074	1.184	.175*	.074	1.192
SF(Mistreatment)	-.494**	.114	.610	-.487**	.114	.614	-.493**	.114	.611	-.493**	.114	.611
SF(Incompetence)	-.224**	.055	.799	-.219**	.055	.803	-.224**	.055	.800	-.223**	.054	.800
RA(Solution Guidance)	-.105**	.036	.900	-.125**	.037	.882	-.106**	.036	.900	-.067	.037	.935
RA(Problem Handling)	-.004	.043	.996	-.010	.043	.990	-.004	.043	.996	.035	.044	1.035
RA(Information/Explanation)	-.088	.047	.916	-.091	.047	.913	-.088	.047	.916	-.091	.047	.913
RA(Channel Directing)	.063	.052	1.065	.049	.053	1.050	.063	.052	1.065	.061	.052	1.063
RA(Follow-Up)	.378**	.042	1.459	.365**	.042	1.440	.378**	.042	1.459	.360**	.042	1.433
RA(Apology)	-.326**	.049	.722	-.340**	.049	.712	-.326**	.049	.722	-.334**	.049	.716
RA(Active Help)	.077	.051	1.080	.084	.051	1.087	.077	.051	1.080	.058	.051	1.060
RA(Showing Empathy)	-.414**	.036	.661	-.435**	.036	.647	-.415**	.036	.660	-.414**	.036	.661
MAc				-.207**	.044	.813						
MAo							-.004	.015	.996			
CR										-.153**	.034	.858
-2 Log likelihood	11819.912			11794.655			11819.816			11798.552		
Cox & Snell R Square	.036			.037			.036			.037		
Nagelkerke R Square	.069			.072			.069			.072		

* p<0.05, ** p< 0.01

Table 3.6 Moderating Effects of Other Twitter Users' Involvement and Customer Co-Recovery

	Model 2a			Model 2b			Model 2c		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Constant	-2.140**	.048	.118	-2.210**	.047	.110	-2.364**	.056	.094
Inter	.307**	.019	1.360	.310**	.019	1.363	.348**	.021	1.416
Soc(Following Number)	.000*	.000	1.000	.000*	.000	1.000	.000*	.000	1.000
SF(Telecoms Issue)	-.227**	.030	.797	-.236**	.030	.790	-.231**	.029	.794
SF(System Malfunction)	.149*	.074	1.161	.180*	.074	1.197	.184*	.074	1.202
SF(Mistreatment)	-.492**	.114	.612	-.463**	.112	.630	-.481**	.112	.618
SF(Incompetence)	-.221**	.055	.802	-.225**	.055	.798	-.217**	.054	.805
RA(Solution Guidance)	-.119**	.038	.888	-.126**	.038	.882	-.008	.048	.992
RA(Problem Handling)	.003	.045	1.003	.030	.046	1.030	.085	.060	1.089
RA(Information/Explanation)	-.102	.049	.903	-.122*	.050	.885	-.057	.063	.945
RA(Channel Directing)	.065	.054	1.067	.091	.056	1.096	.161*	.065	1.174
RA(Follow-Up)	.370**	.043	1.447	.428**	.045	1.534	.538**	.055	1.712
RA(Apology)	-.356**	.051	.701	-.314**	.051	.731	-.205**	.062	.815
RA(Active Help)	.102	.053	1.107	.127*	.055	1.136	.166**	.061	1.180
RA(Showing Empathy)	-.440**	.038	.644	-.440**	.039	.644	-.511**	.045	.600
MAc	-.161*	.067	.851						
MAo				.013	.010	1.013			
CR							.020	.051	1.020
MAc × RA(Solution Guidance)	-.057	.066	.944						
MAc × RA(Problem Handling)	-.057	.064	.945						
MAc × RA(Channel Directing)	-.114	.106	.892						
MAc × RA(Information)	.054	.068	1.056						
MAc× RA(Apology)	.109	.085	1.115						
MAc × RA(Active Help)	-.087	.087	.917						
MAc × RA(Follow-up)	-.035	.072	.966						
MAc × RA(Showing Empathy)	.011	.057	1.011						
MAo × RA(Solution Guidance)				.016	.026	1.016			
MAo × RA(Problem Handling)				-.083*	.038	.920			
MAo × RA(Channel Directing)				-.086	.048	.918			
MAo × RA(Information)				.035	.037	1.036			
MAo × RA(Apology)				-.008	.030	.992			
MAo × RA(Active Help)				-.085	.045	.918			
MAo × RA(Follow-up)				-.088*	.035	.916			
MAo × RA(Showing Empathy)				.043	.028	1.044			
CR × RA(Solution Guidance)							-.036	.020	.964
CR × RA(Problem Handling)							-.043	.035	.958
CR × RA(Channel Directing)							-.103*	.041	.902
CR × RA(Information)							-.037	.040	.964
CR × RA(Apology)							-.135**	.045	.873
CR × RA(Active Help)							-.147**	.051	.863
CR × RA(Follow-up)							-.152**	.034	.859
CR × RA(Showing Empathy)							.087**	.030	1.091
-2 Log likelihood	11788.700			11782.614			11725.617		
Cox & Snell R Square	.037			.038			.041		
Nagelkerke R Square	.073			.073			.079		

* $p < 0.05$, ** $p < 0.01$.

3.5.2 Experiment 2

Interaction frequency (*Inter*) during service recovery was found to be an important factor in Experiment 1. In the same vein with prior research (Choi and Kim 2013; Grönroos 1988), we evaluated interaction quality based on how customers perceived what they had received at each interaction. In Experiment 2 we analysed the service-recovery experience (*SRE*) by modelling interaction quality during service recovery. The logistic regression modelling was again employed in this experiment. We specify Eq. 3:

$$SRE_{tki} = \gamma_0 + \sum_{\substack{k=1\\t=1}} \gamma_{tk} \mathrm{RA}_{tki} + \varepsilon_{tki} \quad , t = 1, \dots, 4; k = 1, \dots 11$$

Eq. 3

In Eq. 3, SRE_{tki} is the emotional status (positive/non-positive emotions) of complainer i when experiencing service-recovery activity k at time t. Notably, in Experiment 1 we set RA_i as a numerical variable, which is the total number of service-recovery activities mentioned in Twitter dialogues. In contrast, we set RA_{tki} as a dummy variable to represent service-recovery activity k at interaction time t, and coded it as 1 if the recovery activity was present at t interaction; otherwise, it was coded as 0. The maximum number of k was 11, which represents the 11 different types of service-recovery activity identified in Phase 1. Notably, we limited t to four (the first four interactions) to prevent the models becoming too large to interpret. Moreover, in our Twitter data, we found that approximately 90 per cent of service-recovery dialogues ended at the fourth interaction. Therefore, it was considered a suitable number.

Experiment 2 Results

Table 3.7 provides the results of the service-recovery activities (*RA*) that affected customer perceived interaction quality in the first four Twitter interactions. Model 3 shows that in the first interaction, the presence of six *RA* variables had an impact on creating customers' positive emotions. These are *Solution Guidance* ($\beta = 0.441$, $p < 0.01$), *Follow-Up* ($\beta = 0.630$, $p < 0.01$), *Problem Handling* ($\beta = 0.572$, $p < 0.01$), *Information/Explanation* ($\beta = 0.149$, $p < 0.01$), *Apology* ($\beta = 0.253$, $p < 0.01$) and *Channel Directing* ($\beta = 0.746$, $p < 0.01$). After the first interaction, only *Follow-Up* was found to have an ongoing positive influence on customer emotions in the following interactions. The presence of *Active Help* influenced customer emotions positively in the second ($\beta = 0.327$, $p < 0.01$) and third interactions ($\beta = 0.612$, $p < 0.01$), although its impact was not supported in the first and fourth interactions. In contrast, two *RA* variables, namely, *Showing Empathy* and *Problem Acknowledgement*, were found to possess a negative influence on customer emotions at each dialogical interaction.

Table 3.7 Results of Recovery Activity Modelling in the First Four Interactions

		Model 3		
		B	S.E.	Exp(B)
t=1	Showing Empathy	-.575**	.035	.563
	Solution Guidance	.441**	.042	1.555
	Follow-Up	.630**	.049	1.877
	Problem Handling	.572**	.058	1.772
	Information/Explanation	.149**	.049	1.161
	Apology	.253**	.051	1.287
	Active Help	.001	.045	1.001
	Channel Directing	.746**	.064	2.108
	Problem Acknowledgement	-.362**	.112	.696
	Guarantee	-.154	.115	.857
	Refund &Replacement	-.183	.352	.833
t=2	Showing Empathy	-.826**	.091	.438
	Solution Guidance	-.428**	.091	.652
	Follow-Up	.732**	.080	2.079
	Problem Handling	-.071	.077	.931
	Information/Explanation	-.280**	.106	.756
	Apology	-.438**	.109	.645
	Active Help	.327**	.115	1.387
	Channel Directing	-.058	.109	.944
	Problem Acknowledgement	-.403*	.200	.668
	Guarantee	-.011	.185	.989
	Refund &Replacement	-.859	.730	.424
t=3	Showing Empathy	-.446**	.130	.640
	Solution Guidance	-.464**	.115	.629
	Follow-Up	.657**	.108	1.929
	Problem Handling	.115	.106	1.122
	Information/Explanation	-.632**	.138	.532
	Apology	-.210	.137	.810
	Active Help	.612**	.171	1.843
	Channel Directing	-.182	.149	.834
	Problem Acknowledgement	-.694**	.260	.499
	Guarantee	-.415	.254	.660
	Refund &Replacement	.198	.487	1.219
t=4	Showing Empathy	-.658**	.187	.518
	Solution Guidance	-.348*	.140	.706
	Follow-Up	.358*	.143	1.431
	Problem Handling	-.258	.159	.773
	Information/Explanation	-.056	.168	.946
	Apology	-.598**	.205	.550
	Active Help	.388	.269	1.473
	Channel Directing	-.320	.203	.726
	Problem Acknowledgement	-1.094**	.397	.335
	Guarantee	-.183	.289	.833
	Refund &Replacement	.238	.634	1.269

* $p < 0.05$, ** $p < 0.01$

3.6 Discussion

3.6.1 Theoretical Implications

Social media service recovery promotes several types of interaction between complainers and other actors, and between complainers and companies, resulting in a dynamic value co-creation process. Figure 3.2 shows the conceptual framework which positions the three dynamic constructs of service recovery based on the Grönroos and Voima's (2013) value creation sphere. Three spheres that promote value creation and co-creation are specified: customer sphere, company sphere and joint sphere (Grönroos and Voima 2013).

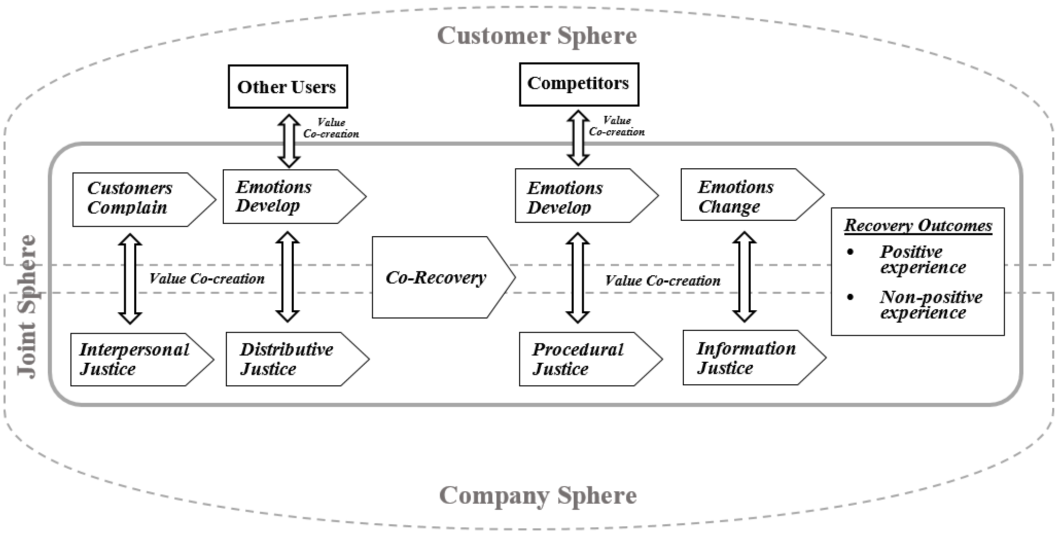

Figure 3.2 Conceptual Framework for Dynamic Service Recovery

Service-recovery dialogues are initiated from the customer complaints carrying customer negative emotions in the customer sphere. These complaints are followed up by the company to restore customers' feelings of justice using the resources regarding distributive, procedural, interpersonal and informational justice in the company sphere. Such dialogues form an interactive process (joint sphere) where both parties can influence each other and co-create or co-destruct value (Grönroos and Voima 2013).

Specifically, this research highlights customer emotions as a key to evaluating interaction quality during service recovery. In the same vein as Heinonen et al. (2010), customer value is regarded as an interactive relativistic preference experience shaped by the emotional perception of companies' resource inputs. Importantly, emotions are embedded in the customers' context, activities and experiences, together with the service company's activities, influencing the customer's future behaviour (Heinonen et al. 2010).

Value-enabling Recovery Offerings in Company Sphere

By examining Twitter dialogs, eleven types of service recovery activities were extracted. Our work reveals critical recovery activities that influence customers' service recovery experience. Firstly, a high level of interpersonal justice offerings, especially negative emotional offerings (showing empathy and apology) delivered by webcare staff, would not promote positive service recovery experience. *Showing Empathy* possesses the highest frequency (71%) out of all company reply tweets. Hocutt et al. (2006) stressed when dissatisfied customers receive empathy from frontline employees, they are inclined to regain a higher level of satisfaction. However, the findings in Phase 2 reveal that showing empathy had a negative impact on service recovery experience. Gilmore and Moreland (2000), studying call centre customer experience, pointed out if the company cannot provide appropriate customer support, showing empathy to customers will not lead to positive customer experiences. Similar results could be found in *Apology*. In terms of positive interpersonal justice offerings, the construct, *Active Help*, was analysed. However, its positive impact was only found in the certain interactions rather than the whole service recovery experience.

The second finding is that procedural justice is the key to positive recovery experience. *Follow-up* encourages the next interaction by asking customers about the progress of service recovery. As for *Channel Directing*, its positive impact is found on the first interaction but not significant for the whole recovery experience. Such an investigation offers companies an understanding of the value co-creation process over the period of service recovery and allows companies to detect the occurrence of value co-destruction (Echeverri and Skalen 2011).

Multi-actor Value Co-creation in Customer Sphere

Social media interactions among complainers and other actors take place in the sphere closed to the company, and thus limited value co-creation occurs between the company and its customers. In the multi-actor engagement cases, we found that competitor involvement had a significant negative impact on the service-recovery experience. Today's competitors are more active on social media, and the virtual presence of competitors can directly lower the threshold of brand switching. Regarding other user involvement, we found no evidence about the main effects, contrary to what has been found in the relevant literature (Schaefers and Schamari 2016). However, our findings pointed out significant negative effects on the relationship between company-initiated recovery activities (*Problem Handling*, *Follow-Up*) and the service-recovery experience when moderated by other users' involvement. This indicates that the potential value co-creation among complaints and multiple actors on social media hinder the company's service-recovery efforts.

Co-creation Recovery in Joint Sphere

Value co-creation between customers and the company takes place when they can influence each other through interactions (Grönroos and Voima 2013). According to Dong et al. (2008), customers can co-create value at many points of the value network within which they obtain the knowledge and skills to perform service recovery. The findings from Phase 1 show that approximately 45 per cent of Twitter dialogues include customer co-created recovery. We identified a variety of co-recovery activities between customers and the company. For example, the company offered informational justice by explaining the reasons for the service problem or informing customers of the problem-handling progress (Lee and Park 2010). Often, informational justice offerings rely heavily on complainer's resource inputting, such as location (e.g. postcode) or their social connections (e.g. other users who have the same problem).

Notably, our findings in Phase 2 show that the increasing engagement of customers in service recovery is detrimental to service recovery. Moreover, customer co-recovery has a negative impact on the relationship between a number of company recovery activities such as *Channel Directing*, *Apology*, *Active Help* and *Follow-Up*, and the service-recovery experience. For example, the main effect of *Follow-Up* has a significant positive impact on the service-recovery experience; yet when moderated by customer co-recovery, *Follow-Up* has a significant negative impact on the service-recovery experience. This finding is in contrast with Dong et al. (2008) and Roggeveen et al. (2012). When complainers are treated as partial employees in terms of conducting service recovery, the company's misevaluation of customers' capabilities and knowledge may result in co-recovery failure. In addition, Roggeveen et al. (2012) explained that customers may view co-creation negatively, as they view it as extra work, which harms customers' evaluation of a company's recovery efforts.

3.6.2 Managerial Implications

Social media acts as a popular organisational tool supporting the functions of customer care, public relations and marketing. Yet, a poor performance of service recovery via social media (11% success rate) is revealed. This finding calls for more attention by service providers to revise and improve the customer experience in the complaint-handling process via social media. Our study is the first field study to investigate real-world service-recovery dialogs between customers and the company using text mining, as opposed to qualitative and experimental studies. Several practical suggestions are given in this research. First, our work proposes an ontology solution to facilitate information extraction and classification following a well-defined typology. The activities, resources and contexts during service recovery are uncovered by the text-mining approach. This approach allows companies to efficiently examine social media webcare, in which dialogical interactions are recorded in a format compatible with big data that facilitates automated analysis and helps to improve organisational resource and activity management when dealing with customer complaints.

The second suggestion is that, since social media dialogs are conducted in a dynamic social environment, data analysis should not be independent of the context. Analysing dialogs on social media enables practitioners to obtain a complete picture of service interactions and to unearth the associative effects of the dynamic components (e.g. multiple participants in dialogs). Current text-mining tools are mostly designed to examine customer opinions rather than complex dialogical structures. This oversimplifies the complicated social contexts and ignores latent factors such as changes in customer emotions. Our work offers a novel approach to exploring social media in terms of how multiple actors influence dialogs and how the company's socially bound reactions co-shape the customer experience. Our approach proposes a way in which text mining could be deployed in a social media analytics context, taking into consideration the mutual influences between dialogue participants.

Finally, we highlight the importance of an interaction-based perspective to analyse the service-recovery experience. In relation to Twitter's service recovery, we found that 90% of mini-cases were completed within four interactions. The research findings also indicate that when the interaction time increases, the odds of a positive service-recovery experience increase. Nevertheless, a high level of complainer engagement on social media is not encouraged, as the failure to rectify problems may drive further customer complaints in the public domain and thus be visible to other customers for a longer period of time. Instead, we suggest that companies need to regularly review the service-recovery interactions based on the change in customer emotions. Monitoring how customer emotions change can offer service providers instant feedback in order to make decisions for the next action. As service recovery is viewed as interaction-driven, the companies' provisions become an environment whereby consumers can create their own unique experience (Prahalad and Ramaswamy 2004). Companies' incompetence in supporting customer value co-creation may lead to the occurrence of value co-destruction (Echeverri and Skalen 2011). Importantly, changes in customer emotions in service recovery could be captured and realised in the customer journey design or service blueprinting. This would help service providers to reduce unfavourable and ineffective service-recovery activities and to promote a positive customer experience.

3.7 Limitations and Future Research

Our work has practical value by utilising text mining and modelling social dialogues to improve customer experience during service recovery. The research methods proposed in this study demonstrate how marketing research can make use of the emerging data analytics techniques to process and extract value from social media big data, thereby contributing to domain knowledge and theory. However, several limitations are highlighted for the avenue of further research. First, the proposed ontology is often restricted to the situation under study (Friman and Edvardsson 2003). Although the ontology development follows previous works of complaint and service-recovery research, the studied situation and the classification scheme are still restricted to the selected industry and social media. Understanding the language structures, user interactions and platform characteristics is essential in developing ontology-based text mining (Ordenes et al. 2014). Twitter webcare, as a micro-blogging service, possesses its own unique communication styles. Hence, the proposed ontology may not be suitable to interpret more diverse social media content and social relationships. However, this research proposes a four-phase process, which could be applied to analyse linguistic-based user interactions on other social media, such as YouTube or Facebook.

Second, we examined the Tweet dialogues based on the observation of visible interactions. If the company's Twitter webcare directs users to other enterprise service channels (e.g. a call centre), some important dialogues may be lost as a result of the inaccessibility of these invisible interactions. In such cases, Twitter dialogues only capture a snapshot of service-recovery interactions, which affects the quality of the research findings. Hence, we suggest that further research should also take multi-channel data into account in analysing service-recovery interactions. Moreover, we believe that our research offers an important step in examining the change in emotions in service-recovery interactions. Together the components – customer emotions, activities, resources, contexts and channels – can offer ample information regarding the customer experience in service recovery. We recommend that future research should extend this research and adopt an overarching view to explore customer journeys by modelling service-recovery interactions.

3.8 References

Aguinis, H. (2004), *Regression Analysis for Categorical Moderators*, Guilford Press.

Ambrose, M., Hess, R.L. and Ganesan, S. (2007), "The Relationship Between Justice and Attitudes: An Examination of Justice Effects on Event and System-related Attitudes," *Organizational Behavior and Human Decision Processes*, 103, 21–36.

Baird, C.H. and Parasnis, G. (2011), "From Social Media to Social Customer Relationship Management," *Strategy & Leadership*, 39(5), 30–37.

Berry, L.L., Wall, E.A. and Carbone, L.P. (2006), "Service Clues and Customer Assessment of the Service Experience: Lessons from Marketing," *Academy of Management Perspectives*, 20(2), 43–57.

Bougie, R., Pieters, R. and Zeelenberg, M. (2003), "Angry Customers Don't Come Back, They Get Back: The Experience and Behavioral Implications of Anger and Dissatisfaction in Services," *Journal of the Academy of Marketing Science*, 31(4), 377–393.

Buttle, F. and Burton, J. (2002), "Does Service Failure Influence Customer Loyalty?" *Journal of Consumer Behaviour*, 1(3), 217–227.

Cao, D., Li, Z. and Ramani, K. (2011), "Ontology-based Customer Preference Modeling for Concept Generation," *Advanced Engineering Informatics*, 25(2), 162–176.

Chebat, J.C. and Slusarczyk, W. (2005), "How Emotions Mediate the Effects of Perceived Justice on Loyalty in Service Recovery Situations: An Empirical Study," *Journal of Business Research*, 58(5), 664–673.

Cheung, F.Y.M. and To, W.M. (2016), "A Customer-Dominant Logic on Service Recovery and Customer Satisfaction," *Management Decision*, 54(10), 2524–2543.

Choi, B.J., and Kim, H.S. (2013), "The Impact of Outcome Quality, Interaction Quality, and Peer-to-Peer Quality on Customer Satisfaction with a Hospital Service", *Managing Service Quality: An International Journal*, 23(3), 188-204

Collins, R. (1981), “On the Microfoundations of Macrosociology,” *American Journal of Sociology*, 86(5), 984–1014.

Crié, D. (2003), “Consumers’ Complaint Behaviour. Taxonomy, Typology and Determinants: Towards a Unified Ontology,” *Journal of Database Marketing & Customer Strategy Management*, 11(1), 60–79.

Davidov, D., Tsur, O. and Rappoport, A. (2010), “Semi-supervised Recognition of Sarcastic Sentences in Twitter and Amazon,” In *Fourteenth Conference on Computational Natural Language Learning*, 107–116.

Davidow, M. (2003), “Have You Heard the Word? The Effect of Word of Mouth on Perceived Justice, Satisfaction and Repurchase Intentions Following Complaint Handling,” *Journal of Consumer Satisfaction, Dissatisfaction & Complaining Behavior*, 16, 66–68.

Dawes, J. (2009), “The Effect of Service Price Increases on Customer Retention,” *Journal of Service Research*, 11(3), 232–245.

Dixon, M., Freeman, K. and Toman, N. (2010), “Stop Trying to Delight Your Customers,” *Harvard Business Review*, 88 (7/8), 116–22.

Dong, B., Evans, K.R. and Zou, S. (2008), “The Effects of Customer Participation in Co-Created Service Recovery,” *Journal of the Academy of Marketing Science*, 36(1), 123–137.

Echeverri, P. and Skalen, P. (2011), “Co-creation and Co-destruction: A Practice-Theory Based Study of Interactive Value Formation,” *Marketing Theory*, 11(3), 351–373.

Edvardsson, B., Tronvoll, B. and Höykinpuro, R. (2011), “Complex Service Recovery Processes: How to Avoid Triple Deviation,” *Managing Service Quality*, 21(4), 331–349.

Einwiller, S.A. and Steilen, S. (2014), "Handling Complaints on Social Network Sites – An Analysis of Complaints and Complaint Responses on Facebook and Twitter Pages of Large US Companies," *Public Relations Review*, 41(2), 195–204.

Ekinci, Y. and Dawes, P.L. (2009), "Consumer Perceptions of Frontline Service Employee Personality Traits, Interaction Quality, and Consumer Satisfaction," *Service Industries Journal*, 29(4), 503–521.

Fan, Y. and Niu, R.H. (2016), "To Tweet or not to Tweet? Exploring the Effectiveness of Service Recovery Strategies Using Social Media." *International Journal of Operations & Production Management*, 36(9), 1014–1036.

Friman, M. and Edvardsson, B. (2003), "A Content Analysis of Complaints and Compliments," *Managing Service Quality*, 13(1), 20–26.

Gallaugher, J. and Ransbotham, S. (2010), "Social Media and Customer Dialog Management at Starbucks," *MIS Quarterly Executive*, 9(4), 197–212.

Gilmore, A. and Moreland, L. (2000), "Call Centres: How can Service Quality be Managed?" *Irish Marketing Review*, 13(1), 3–11.

Goodwin, C. and Ross, I. (1992), "Consumer Responses to Service Failures: Influence of Procedural and Interactional Fairness Perceptions," *Journal of Business Research*, 25(2), 149–163.

Gregg, D.G. and Scott, J.E. (2008), "A Typology of Complaints about eBay Sellers," *Communications of the ACM*, 51(4), 69–74.

Grégoire, Y., Salle, A. and Tripp, T.M. (2014), "Managing Social Media Crises with Your Customers: The Good, the Bad, and the Ugly," *Business Horizons*, 58(2), 173–182.

Grewal, D., Roggeveen, A.L. and Tsiros, M. (2008), "The Effect of Compensation on Repurchase Intentions in Service Recovery," *Journal of Retailing*, 84(4), 424–434.

Grönroos, C. (1984), "A Service Quality Model and its Marketing Implications," *European Journal of Marketing*, 18(4), 36–44.

Grönroos, C. (1988), “Service Quality: The Six Criteria of Good Perceived Service Quality,” *Review of Business*, 9(3), 10–13.

Grönroos, C. (2004), “The Relationship Marketing Process: Communication, Interaction, Dialogue, Value,” *Journal of Business & Industrial Marketing*, 19(2), 99–113.

Grönroos, C. and Voima, P. (2013), “Critical Service Logic: Making Sense of Value Creation and Co-creation.” *Journal of the Academy of Marketing Science*, 41(2), 133–150.

Gruber, T.R. (1993), “A Translation Approach to Portable Ontology Specifications,” *Knowledge Acquisition*, 5(2), 199–220.

Haghighi, P.D., Burstein, F., Zaslavsky, A. and Arbon, P. (2013), “Development and Evaluation of Ontology for Intelligent Decision Support in Medical Emergency Management for Mass Gatherings.” *Decision Support Systems*, 54(2), 1192–1204.

Heinonen, K., Tore Strandvik, T., Mickelsson, K.J., Edvardsson, B., Sundström, E. and Andersson, P. (2010), “A Customer-Dominant Logic of Service,” *Journal of Service Management*, 21(4), 531–548.

Hocutt, M.A., Bowers, M.R. and Donavan, D.T. (2006), “The Art of Service Recovery: Fact or Fiction? *Journal of Services Marketing*, 20(3), 199–207.

Hosmer, D.W., Lemeshow, S. and Sturdivant, R.X. (2013), *Applied Logistic Regression* 3rd Eds., John Wiley & Sons.

Huang, W. (2008), “The Impact of Other-Customer Failure on Service Satisfaction,” *International Journal of Service Industry Management*, 19(4), 521–536.

Kaplan, A.M. and Haenlein, M. (2010), “Users of the World, Unite! The Challenges and Opportunities of Social Media,” *Business Horizons*, 53(1), 59–68.

King, J.E. (2008), “Binary Logistic Regression,” In J.W. Osborne, ed. *Best Practices in*

Quantitative Methods. Sage Publications.

Kozinets, R.V. (2002), "The Field behind the Screen: Using Netnography for Marketing Research in Online Communities," *Journal of Marketing Research*, 39(1), 61–72.

Kunz, W.H., Aksoy, L., Bart, Y., Heinonen, K.K.S., Ordenes, F.V., Sigala, M., Diaz, D. and Theodoulidis, B. (2017), "Customer Engagement in a Big Data World," *Journal of Services Marketing*, 31(2), 161–171.

Larivière, B., Joosten, H., Malthouse, E.C., van Birgelen, M., Aksoy, P., Kunz, W.H. and Huang, M.- H. (2013), "Value Fusion: The Blending of Consumer and Firm Value in the Distinct Context of Mobile Technology and Social Media," *Journal of Service Management*, 24(3), 268–293.

Lee, C.-H., Wang, Y.-H. and Trappey, A.J.C. (2015), "Ontology-based Reasoning for the Intelligent Handling of Customer Complaints," *Computers & Industrial Engineering*, 84, 144–155.

Lee, E. and Park, J. (2010), "Service Failures in Online Double Deviation Scenarios: Justice Theory Approach," *Managing Service Quality: An International Journal*, 20(1), 46–69.

Lemon, K.N. and Verhoef, P.C. (2016), "Understanding Customer Experience and the Customer Journey," *Journal of Marketing*, 80(6), 69–96.

Liau, B.Y. and Tan, P.P. (2014), "Gaining Customer Knowledge in Low Cost Airlines through Text Mining," *Industrial Management & Data Systems*, 114(9), 1344–1359.

Libai, B., Bolton, R., Bügel, M.S., de Ruyter, K., Götz, O., Risselada, H., and Stephen, A.T. (2010), "Customer-to-Customer Interactions: Broadening The Scope of Word of Mouth Research," *Journal of Service Research*, 13(3), 267–282.

Magnini, V.P., Ford, J. B., Markowski, E.D., and Honeycutt, E.D.J. (2007), "The Service Recovery Paradox: Justifiable Theory or Smoldering Myth?" *Journal of Services*

Marketing, 21(3), 213–225.

Manning, C.D., Raghavan, P. and Schütze, H. (2008), *Introduction to Information Retrieval*, Cambridge: Cambridge University Press.

Maxham, J.G. (2001), "Service Recovery's Influence on Consumer Satisfaction, Positive Word-of-Mouth, and Purchase Intentions," *Journal of Business Research*, 54(1), 11–24.

Meuter, M.L. and Bitner, M.J. (1998), "Self-service Technologies: Extending Service Frameworks and Identifying Issues for Research," In *AMA Winter Educator's Conference Proceedings: Marketing Theory and Applications*, 12–19.

Mikroyannidis, A. (2007), Toward a Social Semantic Web. *Computer*, 40(11), 113–115.

Mikroyannidis, A. and Theodoulidis, B. (2010), "Ontology Management and Evolution for Business Intelligence," *International Journal of Information Management*, 30(6), 559–566.

Missikoff, M., Navigli, R. and Velardi, P. (2002), "The Usable Ontology: An Environment for Building and Assessing a Domain Ontology," In *The Semantic Web—ISWC 2002*, 39–53.

Neslin, S.A., Gupta, S. Kamakura, W., Lu, J. and Mason, C.H. (2006), "Defection Detection: Measuring and Understanding the Predictive Accuracy of Customer Churn Models," Journal of Marketing Research, 43(2), 204–211.

Nicholls, R. (2010), "New Directionsfor Customer-to-Customer Interaction Research," *Journal of Services Marketing*, 24(1), 87–97.

Novani, S. and Kijima, K. (2012), "Value Co-creation by Customer-to-Customer Communication: Social Media and Face-to-Face for Case of Airline Service Selection," *Journal of Service Science and Management*, 5(1),101–109.

Ordenes, F. V, Theodoulidis, B., Burton, J., Gruber, T. and Zaki, M. (2014), "Analyzing Customer Experience Feedback Using Text Mining: A Linguistics-Based Approach," *Journal of Service Research*, 17(3), 278–295.

Ozgen, O. and Kurt, S.D. (2012), "Pre-recovery and Post-recovery Emotions in the Service Context: A Preliminary Study," *Managing Service Quality*, 22(6), 592–605.

Pfeffer, J., Zorbach, T. and Carley, K.M. (2014), "Understanding Online Firestorms: Negative Word-of-Mouth Dynamics in Social Media Networks," *Journal of Marketing Communications*, 20(1–2), 117–128.

Prahalad, C.K. and Ramaswamy, V. (2004), "Co-creation Experiences: The Next Practice in Value Creation. *Journal of Interactive Marketing*, 18(3), 5–14.

Riloff, E., Qadir, A., Surve, P., De Silva, L., Gilbert, N. and Huang, R. (2013), "Sarcasm as Contrast between a Positive Sentiment and Negative Situation," In *Empirical Methods in Natural Language Processing (EMNLP 2013)*, 704–714.

del Río-Lanza, A.B., Vázquez-Casielles, R. and Díaz-Martín, A.M. (2009), "Satisfaction with Service Recovery: Perceived Justice and Emotional Responses," *Journal of Business Research*, 62(8), 775–781.

Roggeveen, A.L., Tsiros, M. and Grewal, D. (2012), "Understanding the Co-creation Effect: When Does Collaborating with Customers Provide A Lift to Service Recovery?" *Journal of the Academy of Marketing Science*, 40(6), 771–790.

Schaefers, T. and Schamari, J. (2016), "Service Recovery via Social Media: The Social Influence Effects of Virtual Presence," *Journal of Service Research*. 19(2), 192–208.

Schoefer, K. and Ennew, C. (2005), "The Impact of Perceived Justice on Consumers' Emotional Responses to Service Complaint Experiences," *Journal of Services Marketing*, 19(5), 261–270.

Sharma, S., Durand, R.M. and Gur-Arie, O. (1981), "Identification and Analysis of

Moderator Variables," *Journal of Marketing Research*, 18(3), 291–300.

Smith, A.K., Bolton, R.N. and Wagner, J. (1999), "A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery," *Journal of Marketing Research*, 36(3), 356–372.

Smith, A.K. and Bolton, R.N. (2002), "The Effect of Customers' Emotional Responses to Service Failures on Their Recovery Effort Evaluations and Satisfaction Judgments," *Journal of the Academy of Marketing Science*, 30(1), 5–23.

Solis, B. and Kutcher, A. (2011), *Engage! The Complete Guide for Brands and Businesses to Build, Cultivate, and Measure Success in the New Web, Revised and Updated*, John Wiley & Sons.

Stein, A. and Ramaseshan, B. (2015), "Customer Referral Behavior: Do Switchers and Stayers Differ?" *Journal of Service Research*, 18(2), 229–239.

Tax, S.S., Brown, S.W., and Chandrashekaran, M. (1998), "Customer Evaluations of Service Complaint Experiences: Implications for Relationship Marketing," *Journal of Marketing*, 7(3), 3–30.

Tripp, T.M. and Grégoire, Y. (2011), "When Unhappy Customers Strike Back on the Internet," *MITSloan Management Review*, 52(3), 37–44.

Tronvoll, B. (2011), "Negative Emotions and their Effect on Customer Complaint Behaviour," *Journal of Service Management*, 22(1), 111–134.

Verhoef, P.C., Lemon, K.N., Parasuraman, A., Roggeveen, A., Tsiros, M. and Schlesinger, L.A. (2009), "Customer Experience Creation: Determinants, Dynamics and Management Strategies," *Journal of Retailing*, 85(1), 31–41.

Willemsen, L., Neijens, P.C. and Bronner, F.A. (2013), "Webcare as Customer Relationship and Reputation Management? Motives for Negative Electronic Word of

Mouth and Their Effect on Webcare Receptiveness," In S. Rosengren, M. Dahlén, and S. Okazaki, eds. *Advances in Advertising Research (Vol. IV)*. EAA Series. Wiesbaden: Springer Fachmedien Wiesbaden, 55–69.

Xu, K., Liao, S.S., Li, J., and Song, Y. (2011), "Mining Comparative Opinions from Customer Reviews for Competitive Intelligence," *Decision Support Systems*, 50(4), 743–754.

Xu, Y., Marshall, R., Edvardsson, B and Tronvoll, B. (2014), "Show You Care: Initiating Co-creation in Service Recovery," *Journal of Service Management*, 25(3), 369–387.

Appendix A:

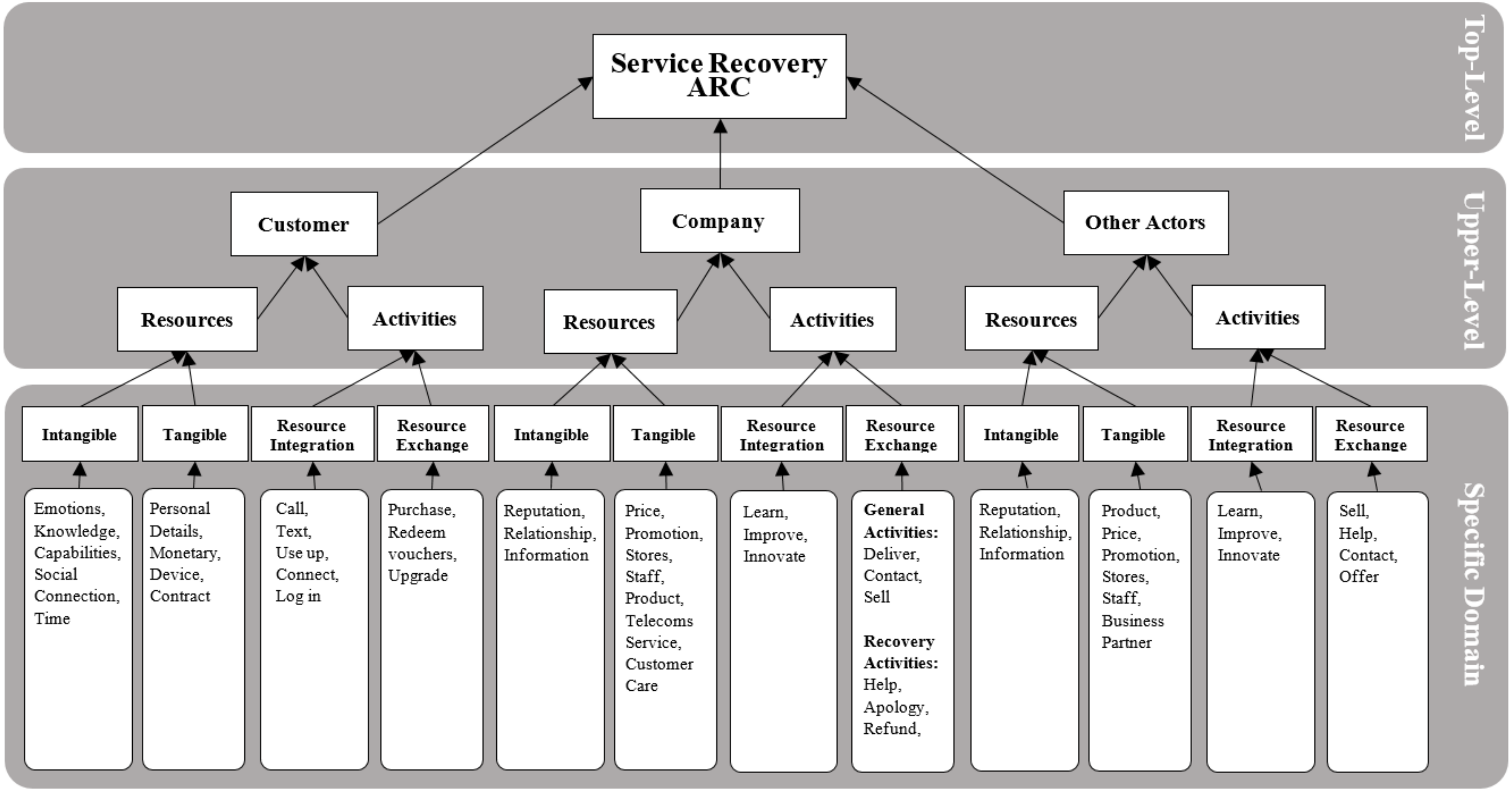

Figure 3.3 The Ontologies for Information Extraction

Appendix B:

Table 3.8 Descriptive Statistics of Predictor Variables

Total Cases: 17,125				
	Min.	Max.	Mean	Std Deviation
Soc(Follower Number)	0	1226122	889.54	11252.688
Soc(Following Number)	0	130524	589.22	2125.594
Inter	1	32	2.38	1.767
MAc	0	28	0.23	0.718
MAo	0	190	0.46	1.978
CR	0	10	0.65	0.929
SF(Telecoms Issue)	0	92	0.78	1.225
SF(Incompetence)	0	21	0.2	0.548
SF(Contact Company)	0	14	0.17	0.502
SF(Wait Time)	0	8	0.13	0.421
SF(System Malfunction)	0	5	0.08	0.302
SF(Price and Payment)	0	11	0.07	0.331
SF(Mistreatment)	0	8	0.06	0.313
SF(Mistake)	0	4	0.05	0.262
SF(Delivery and Not Receiving)	0	4	0.03	0.199
SF(Product and Stock)	0	6	0.03	0.176
SF(Promotion and CRM)	0	5	0.03	0.171
SF(Customer-Initiated Failure)	0	3	0.02	0.145
SF(Fraud)	0	2	0	0.072
SF(Other Companies' Failures)	0	6	0	0.067
RA(Showing Empathy)	0	7	0.92	0.774
RA(Solution Guidance)	0	11	0.53	0.774
RA(Problem Handling)	0	7	0.33	0.62
RA(Follow-Up)	0	7	0.31	0.557
RA(Information/Explanation)	0	5	0.3	0.566
RA(Apology)	0	7	0.29	0.564
RA(Active Help)	0	6	0.24	0.486
RA(Channel Directing)	0	5	0.2	0.47
RA(Problem Acknowledgement)	0	4	0.08	0.314
RA(Guarantee)	0	4	0.07	0.296
RA(Refund & Replacement)	0	3	0.01	0.112

Appendix C:

Table 3.9 Pearson Correlation Results between Predictor Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Inter	1	-.004	.280**	.133**	.143**	.300**	.460**	.364**	.333**	.293**	.301**	.304**	.144**	.416**	.070**	.101**	.617**
2. Soc(Following Number)	-.004	1	.001	-.014	.002	-.013	-.016*	.001	.004	-.010	-.012	-.013	-.011	-.041**	.024**	.021**	-.007
3. SF(Telecoms Issue)	.280**	.001	1	.122**	.031**	.247**	.269**	.222**	.249**	-.042**	.236**	.058**	-.103**	.097**	.228**	.462**	.228**
4. SF(System Malfunction)	.133**	-.014	.122**	1	-.003	.076**	.114**	.043**	.052**	.056**	.031**	.038**	-.037**	.041**	-.002	.085**	.101**
5. SF(Mistreatment)	.143**	.002	.031**	-.003	1	.190**	-.032**	.107**	.011	.042**	.009	.106**	.117**	.142**	.051**	.138**	.075**
6. SF(Incompetence)	.300**	-.013	.247**	.076**	.190**	1	.078**	.123**	.053**	.133**	.079**	.145**	.139**	.234**	.119**	.259**	.174**
7. RA(Solution Guidance)	.460**	-.016*	.269**	.114**	-.032**	.078**	1	.052**	.149**	.051**	.269**	-.013	-.142**	.188**	-.048**	.010	.421**
8. RA(Problem Handling)	.364**	.001	.222**	.043**	.107**	.123**	.052**	1	.319**	-.068**	.239**	.106**	.058**	.152**	.012	.041**	.371**
9. RA(Information)	.333**	.004	.249**	.052**	.011	.053**	.149**	.319**	1	-.004	.138**	.115**	-.077**	.023**	.036**	.040**	.246**
10. RA(Channel Directing)	.293**	-.010	-.042**	.056**	.042**	.133**	.051**	-.068**	-.004	1	-.021**	.098**	.039**	.192**	-.032**	-.005	.126**
11. RA(Follow-Up)	.301**	-.012	.236**	.031**	.009	.079**	.269**	.239**	.138**	-.021**	1	.044**	-.001	.129**	-.020**	.065**	.162**
12. RA(Apology)	.304**	-.013	.058**	.038**	.106**	.145**	-.013	.106**	.115**	.098**	.044**	1	.063**	.158**	.003	.043**	.122**
13. RA(Active Help)	.144**	-.011	-.103**	-.037**	.117**	.139**	-.142**	.058**	-.077**	.039**	-.001	.063**	1	.158**	.031**	.024**	-.010
14. RA(Showing Empathy)	.416**	-.041**	.097**	.041**	.142**	.234**	.188**	.152**	.023**	.192**	.129**	.158**	.158**	1	-.054**	-.011	.245**
15. MAc	.070**	.024**	.228**	-.002	.051**	.119**	-.048**	.012	.036**	-.032**	-.020**	.003	.031**	-.054**	1	.311**	-.008
16. MAo	.101**	.021**	.462**	.085**	.138**	.259**	.010	.041**	.040**	-.005	.065**	.043**	.024**	-.011	.311**	1	.026**
17. CR	.617**	-.007	.228**	.101**	.075**	.174**	.421**	.371**	.246**	.126**	.162**	.122**	-0.01	.245**	-.008	.026**	1

* correlation is significant at the 0.05 level (2-tailed), ** correlation is significant at the 0.01 level (2-tailed).

Appendix D:

Table 3.10 Effect Size based on Standardized Features

	Cohen's d	Effect-size r
Inter	.341	.168
Soc (Following Number)	.055	.027
SF (Telecoms Issue)	-.095	-.047
SF (System Malfunction)	.083	.041
SF (Mistreatment)	-.104	-.052
SF (Incompetence)	-.062	-.031
RA (Solution Guidance)	.114	.057
RA (Problem Handling)	.097	.049
RA (Information)	.063	.032
RA (Channel Directing)	.112	.056
RA (Follow-Up)	.266	.132
RA (Apology)	-.080	.040
RA (Active Help)	.074	.037
RA (Showing Empathy)	-.159	-.080
MAc	-.092	-.046
MAo	.008	.004
CR	.092	.046

Chapter 4 A Dialogue-Mining Framework for Improving Service Recovery: Customer Care on Twitter

ABSTRACT

Analysing dialogues generated during complaint handling can help companies to improve internal process recovery and customer recovery. Previous research adopted qualitative analysis approaches to investigate service recovery dialogues yet could only examine a limited volume of data in a particular period. Also, the understanding of dialogues was constrained to the linguistic dimension, with researchers and practitioners examining specific aspects of service recovery that they believe to be important, rather than those aspects representing complainers' real service experiences. To address these gaps, this paper introduces a dialogue-mining framework, specifying three dimensions of dialogues to be thoroughly investigated: linguistic and semantic, process, and relationship. The three dimensions are translated into a dialogue-mining apparatus that facilitates text mining and process mining to extract embedded information regarding customer experience processes and service recovery processes from dialogues. We demonstrate this analytical approach by analysing the company-to-customer dialogues in the context of service recovery, which requires high levels of engagement of both parties to tackle service problems and improve service practices. The outcomes of this research shed light on the theorisation of dialogue analysis methods in marketing research and provide practical value for managing complaint-handling dialogues to avoid negative customer outcomes and utilise complaints to improve practice.

Keywords: Social Media, Service Recovery, Text Mining, Process Mining, Marketing Analytics

4.1 Introduction

For over four decades, research on customer-complaint handling has made insightful contributions in the marketing area. The actions taken to deal with customer complaints and to recover customer satisfaction are conceptualised as service recovery (Grönroos 1988). Effective service recovery is found to possess positive impacts on customer post-recovery satisfaction, repurchase intention and customer relationships (Grewal et al. 2008; Maxham 2001; Smith et al. 1999; Vázquez-Casielles et al. 2010). However, van Vaerenbergh and Orsingher (2016) highlighted that the level of customer satisfaction with service recovery in 2013 was no higher than that reported in 1976.

One of the most important reasons for such a low level of customer post-recovery satisfaction is that companies fail to make use of customer complaints to improve their services/processes and prevent similar failures in the future (van Vaerenbergh and Orsingher 2016). Another potential reason is the change in customer complaining behaviour driven by the advance of digital platforms, especially negative word of mouth on social media (Einwiller and Steilen 2014; Grégoire et al. 2014). Increasingly, companies provide *webcare* to proactively post messages without complainants' requests or to reactively respond to complainants (van Noort and Willemsen 2012). Despite its timeliness in responding to complaints, webcare may lead to negative outcomes when companies fail to deliver appropriate complaint handling in a very public arena (van Noort and Willemsen 2012).

Often, practitioners and researchers alike conduct surveys or interviews to evaluate service recovery performance, yet these methods are often hindered by low response rates (e.g. Tax et al. 1998) or limited sample sizes (e.g. Johnston and Michel 2008). Moreover, prior research tended to focus on the aspects of service recovery performance that the focal company believes to be important, rather than those aspects of the customers' experience that the customers feel are significant (Ordenes et al. 2014). To obtain more nuanced data on a greater scale, some researchers examine service recovery incidents within service dialogues (Fan and Niu 2016; Rafaeli et al. 2008). In most service recovery procedures, human-to-human (or human-to-machine) dialogues – whether taking place in physical stores, through call centres or on online platforms – play a vital role in implementing customer services. Traditionally, researchers analyse dialogue data using qualitative discourse analysis, such as open coding employed by Rafaeli et al. (2008). However, qualitative coding is time-consuming, and researchers are unlikely to comb through a very large volume of dialogue data. For example, Fan and Niu (2016) observed service recovery dialogues between several airline companies and their customers on Twitter during a five-month period, examining only 347 pieces of dialogue.

Researchers increasingly advocate the use of social media analytics and text mining to aid in collecting, analysing and presenting the findings from social media dialogues and to improve business performance (Zeng et al. 2010). Nevertheless, most of the prior research focused exclusively on the dialogue content, or more specifically, the linguistic aspect of dialogues (Fan and Niu 2016; Hwong et al. 2017; Ordenes et al. 2017). Dialogues are a rich source of information containing interactive processes within which participants build up a shared meaning through "acts" (the smallest unit of service activity), "episodes" (a service context that includes a series of acts), "sequences" (interrelated episodes), and "relationship" (the feelings and perceptions developed from the sequences) (Grönroos 2004). The linguistic aspect can only capture "acts" and "episodes" of dialogues, as it considers the natural language features of actor conversations (Ordenes et al. 2014). To our knowledge, none of the existing models, frameworks or analytical processes is designed to understand the "sequences" and "relationship" aspects of dialogues and associate holistic meanings embedded within dialogue interactions.

This paper addresses the research gaps and contributes in terms of theorising dialogue analytical methods and improving service recovery operations via social media interactions. Firstly, this paper introduces a dialogue-mining framework consisting of three dimensions: linguistic and semantic, process, and relationship. These dimensions specify how the embedded "acts", "episodes" (linguistic and semantic dimension), and "sequences" (process dimension) impact the third dimension – customer relationship. Therefore, the framework improves the linguistic-based analysis, allowing for the investigation of dialogue data at different granularities.

Secondly, the utility of the proposed framework is demonstrated by testing it in a real-world setting. Text mining and process mining were used to extract the embedded information related to the three dimensions in the service recovery dialogues and identify the weak links during service recovery. The applications build up a linkage between the marketing theory and the analytical methods, thus assuring the generalisability of the framework. Furthermore, the framework can serve as a problem-solving solution for addressing the identified service recovery issues: process recovery and customer recovery. As a result, it contributes to improving the management of complaint-handling dialogues to avoid negative outcomes on social media.

4.2 Overview of Service Recovery

Tackling customer complaints has long featured as a form of defensive marketing, viewing them as a second chance to recover service failures and avoid customer churn (Fornell and Wernerfelt 1987; Hirschmann 1970). The fundamental premise of service recovery is that since service failure events trigger customers' feelings of injustice, it is only when companies restore customers' perceived justice that they can recover customer satisfaction (Hart et al. 1990; Smith et al. 1999). An extensive amount of work has been conducted to examine how perceived justice influences customer post-recovery satisfaction, repurchase intention and positive word-of-mouth (Buttle and Burton 2002;Davidow 2003; Grewal et al. 2008; Maxham 2001).

In addition to customer recovery, service recovery has a greater impact on the improvement of companies' operational practice regarding process recovery and employee recovery (Johnston and Michel 2008; Michel et al. 2009). Process recovery makes use of customer complaints to improve service delivery systems and re-examines the service quality (Simons and Kraus 2005). Through process improvement, companies can reduce costs by removing the root causes of service problems and avoiding future service failures (Johnston and Michel 2008). Notably, process recovery is not only a "damage-control" tool but also a long-term business strategy that a company devotes to improving inefficient and ineffective processes and driving innovation of service offerings (La and Kandampully 2004).

Employee recovery is a company's internal service recovery (Bowen and Johnston 1999). It focuses on addressing frontline employees' stress and negative emotions during and after their interactions with complainers (Johnston and Michel 2008; Michel et al. 2009). Frontline employees tend to develop feelings of helplessness and alienation, leading to passive and maladapted behaviour and affecting their performance in complaint handling (Bowen and Johnston 1999). Tackling the negative impact of complaint handling on employees allows companies to reduce staff absenteeism and turnover (Johnston and Michel 2008).

Despite numerous research contributions, today's companies are still struggling to manage particular dimensions of service recovery (van Vaerenbergh and Orsingher 2016). This paper especially focuses on the dimensions of process recovery and customer recovery to examine the relevant issues.

Process Recovery

Extant research has taken a variety of approaches to the issues of process recovery, including profiling of service failures, analysing failure types and impacts, and developing operational frameworks for service recovery (Johnston and Michel 2008). Yet, the considerable body of research has had a limited influence on practical process recovery (Michel et al. 2009; van Vaerenbergh and Orsingher 2016). Research conducted by van Vaerenbergh and Orsingher (2016) pointed out that many companies fail to take advantage of customer complaints to improve their service processes. Customer complaints are important data providing insights into process recovery. Compared to approaches such as total quality management (TQM), mystery shopping and the critical incident technique (CIT) for obtaining service failure data (Johnston and Michel 2008), customer feedback provides easy access for companies to understand customers' experiences at any point of a service process (Ordenes et al. 2014).

According to Gentile et al. (2007, p. 397), "the customer experience originates from a set of interactions between a customer and a product, a company, or part of its organisation, which provoke a reaction". During service interactions, customers generate cognitive and emotional responses towards companies' offerings (also known as touchpoints) and make an overall judgement for their service experience (Gentile et al. 2007; Lemke et al. 2010; Lemon and Verhoef 2016). Prior research viewed the customer experience perspective as a useful approach to evaluate service qualities because customers often go beyond the service processes and perceive service value differently (Lemke et al. 2010; Lemon and Verhoef 2016; Ordenes et al. 2014). Nevertheless, applying the customer experience perspective to service recovery research remains under-explored.

Customer Recovery

Customer recovery is becoming evermore complex. The advent of social media fuels customer complaints and creates diverse complaining behaviours (Grégoire et al. 2014). Companies increasingly employ social media webcare to detect and react to customer complaints (Larivière et al. 2013). Some important operational issues were raised by Fan and Niu (2016), such as how to identify key variables as well as their interrelationships for improving the performance of customer recovery, and how to integrate social media webcare with existing customer care platforms.

In contrast to service recovery on company-owned platforms (e.g. call centre, face-to-face problem handling), customer recovery on social media is constrained to the platform's policies, language use, and communication patterns (Abney et al. 2017). Fan and Niu (2016) analysed customer recovery dialogues on Twitter, pointing out that service agents' responses related to certain service recovery offerings have an evident impact on customer emotions and customer satisfaction. Considering that today's customers feel more time pressure to resolve service problems (Grainer et al. 2014) and that inefficient service recovery may lead to complainer exit and other negative effects (van Noort and Willemsen 2012), companies need to understand what and how critical service recovery offerings are delivered via social media and evaluate the corresponding customer outcomes.

Furthermore, social media customer recovery takes place in the public arena, where other users can actively or passively participate (Schaefers and Schamari 2016). Often, companies need to help complainers migrate amongst service channels to tackle problems (Dalla Pozza 2014). Such channel migration may lead to complainers' annoyance as companies "ping-pong" customer complaints between different service touchpoints (Grainer et al. 2014). Prior research exploring channel management issues has shed light on complainant channel choice (Dalla Pozza 2014; Mattila and Wirtz 2004), yet scarce research explores the issue with respect to integrating new channels, such as social media webcare, with traditional channels and improving resource allocation amongst these channels.

Based on the identified obstacles, this research aims to provide a better solution built on the investigation of social media dialogues generated during service recovery. First, this work adds to the understanding of process recovery and suggests a method allowing practitioners to efficiently capture service problems outlined in descriptions of customer experience. Second, a process-centric approach is developed to analyse the sequences of service recovery activities (temporal process) and the associations between recovery activities and service channels (situational process). In this way, we make recommendations for improving customer recovery management. To achieve the research aim, a dialogue-mining framework is proposed in this study.

4.3 Dialogue-Mining Framework

Prior marketing research analysing dialogue data was centred on a methodological level, such as the application of discourse analysis (Rafaeli et al. 2008) or text analytics (Tirunillai and Tellis 2014). For example, Rafaeli et al. (2008) analysed 166 call centre transcripts of a retail bank, identifying five types of employees' customer orientation behaviours associated with customer assistance. The lighting company Osram Sylvania conducted dialogue analysis on telephone service transcripts and found that the words "can't", "won't", and "don't", when used by frontline employees, tended to trigger negative reactions from customers and cause repeated phone calls (Dixon et al. 2010).

In addition to the understanding of analysing dialogues as a research method, dialogue has been viewed as a key component of co-creating value between a company and its customers (Finne and Grönroos 2017; Grönroos 2004; Prahalad and Ramaswamy 2004). Dialogue requires deep engagement, and it builds on the access and transparency to information (Prahalad and Ramaswamy 2004). During interactive dialogues, two parties can learn together by sharing opinions and building mutual value through better understanding of each other's mental models, logics, cognitive constructs, and schemas (Ballantyne 2004; Grönroos 2004). Ballantyne (2004) highlighted that dialogue is key to develop and maintain a relationship as it enables dialogical parties to access to relationship specific knowledge that is built on the past value co-creation experiences and is constantly updated by new interactions.

On the basis of value co-creation literature, we codify a dialogue-mining framework in three operational dimensions: linguistic and semantic, process, and relationship.

Linguistic and Semantic Dimension

Grönroos (2004), building on Holmlund's (1997) work, divided communicative interactions into four hierarchical levels: acts, episodes, sequences, and relationship. Acts are the basic unit of interactions, including any types of interaction modes (e.g. physical goods, services, financial aspects or social contacts), while an episode specifies the context of a flow of acts (e.g. a visit to a telecoms company's local store to discuss a contract). The linguistic and semantic dimension is useful for capturing the "acts" and "episodes" within interactive dialogues.

Linguistic features include length of texts, position of words, and grammar, while semantic features are related to the meaning or sentiment polarity of words (Ordenes et al. 2014). Analysing linguistic and semantic features of dialogue has been widely used by researchers (e.g. Fan and Niu 2016; Rafaeli et al. 2008) to understand the perspectives and motivations of both parties and the crucial elements driving the information exchange. For the service recovery dialogues, Fan and Niu (2016) analysed the speeches of the service agents to reveal how companies encode a series of acts such as apology and compensation and also examined speeches of complainers regarding their emotional responses and satisfaction. Gummesson and Mele (2010) highlighted that interactions are a resource transfer process in which knowledge, products, services, and solutions are exchanged amongst actors, and these resources allow actors to create value for their goals. Company speeches carry the "acts" and "episodes" regarding the company's resources and solutions that a customer can act on, while the real value is realised in customer speeches (Finne and Grönroos 2017).

Process Dimension

The aggregation of “episodes” forms a sequence (Grönroos 2004). The process dimension is important for understanding the interactive “sequences” of value co-creation, which are driven by “acts” and “episodes” in the linguistic and semantic dimension. Gummesson and Mele (2010) noted that participants set up a dialogue process not only to exchange information but also to make their resources available for creating new value that better matches their aims, knowledge, and skills.

Finne and Grönroos (2009, 2017) clarified two aspects of the process that affect the value of communication: time frame and situational context. The time frame aspect depicts the interactions of past, present and future experiences (Finne and Grönroos 2017). For instance, during service processes, customers evaluate the service provider’s performance at each interaction point over time, and the past interactions tend to affect their current interaction and willingness to engage in future interactions (Lemon and Verhoef 2016). The situational aspect incorporates a wide range of components, including internal factors (e.g. actors’ mental status, motivations, and capabilities) and external factors (e.g. connections of associate and competitor activities) (Finne and Grönroos 2017). Since companies are often unable to control dialogue interactions, especially when dialogues happen in the public arena (e.g. social media), they should consider to what extent these two aspects influence the value of communication and how they can manage them.

Relationship Dimension

Growing a “relationship” based on the “sequences” of “episodes” and “acts” is the long-term purpose of dialogue interactions (Grönroos 2004). Relationships are always present wherever interactions occur between two or more actors, yet the quality of relationships is determined based on the experiences of actor interaction over time

(Ballantyne and Varey 2006). Ballantyne (2004) used the term *relationship specific knowledge* to describe the knowledge about how dialogue parties deal with one another and resolve dilemmas to satisfy expectations. Such knowledge is developed from the co-created experiences amongst dialogue parties and affect the future dialogical interaction (Ballantyne 2004; Payne et al. 2008).

Payne et al. (2008) pointed out that the relationship experiences promote customer learning, and the results of learning are manifested in the change of customer attitudes and preference. In other words, customer satisfaction and the degree of customer engagement determine if the relationship is continued (Payne et al. 2008). The role of the company during such relationship experience is to providing experiential encounters, through which customers co-create value by utilising their resources (Prahalad and Ramaswamy 2004). On the other hand, relationship experiences also lead to organisational learning, it helps companies to conduct co-creation and relationship experience design (Payne et al. 2008). More specifically, the relationship dimension allows companies to evaluate their value propositions by observing customers' "value-in-use" and thus, develop customer preferred offerings.

4.4 The Dialogue-Mining Methodology

Our objective is to explore dialogues of social media complaint handling for improving issues regarding process recovery and customer recovery. To do so, a dialogue-mining apparatus was developed following the proposed framework, with three dimensions: linguistic and semantic, process, and relationship. Figure 4.1

presents the dialogue-mining apparatus and its four phases: dialogue data collection, information extraction, process modelling, and relationship evaluation.

4.4.1 Dialogue Data Collection

Dialogue data commonly exists within enterprise systems, such as call centres, email services, live chats, or helpdesk systems. Also, companies increasingly conduct social listening and collect social media dialogues to identify crucial customer insights (Schweidel and Moe 2014). Fan and Niu (2016) viewed service recovery dialogues as “mini cases” that record the back-and-forth conversations between service agents and customers until service problems are resolved or no further responses are given by either side. To mine dialogue data, we suggest that it should contain four elements: (i) case ID, (ii) at least two identifiable participants, (iii) the timestamp of each speech, and (iv) dialogue content (e.g. texts, pictures, or videos).

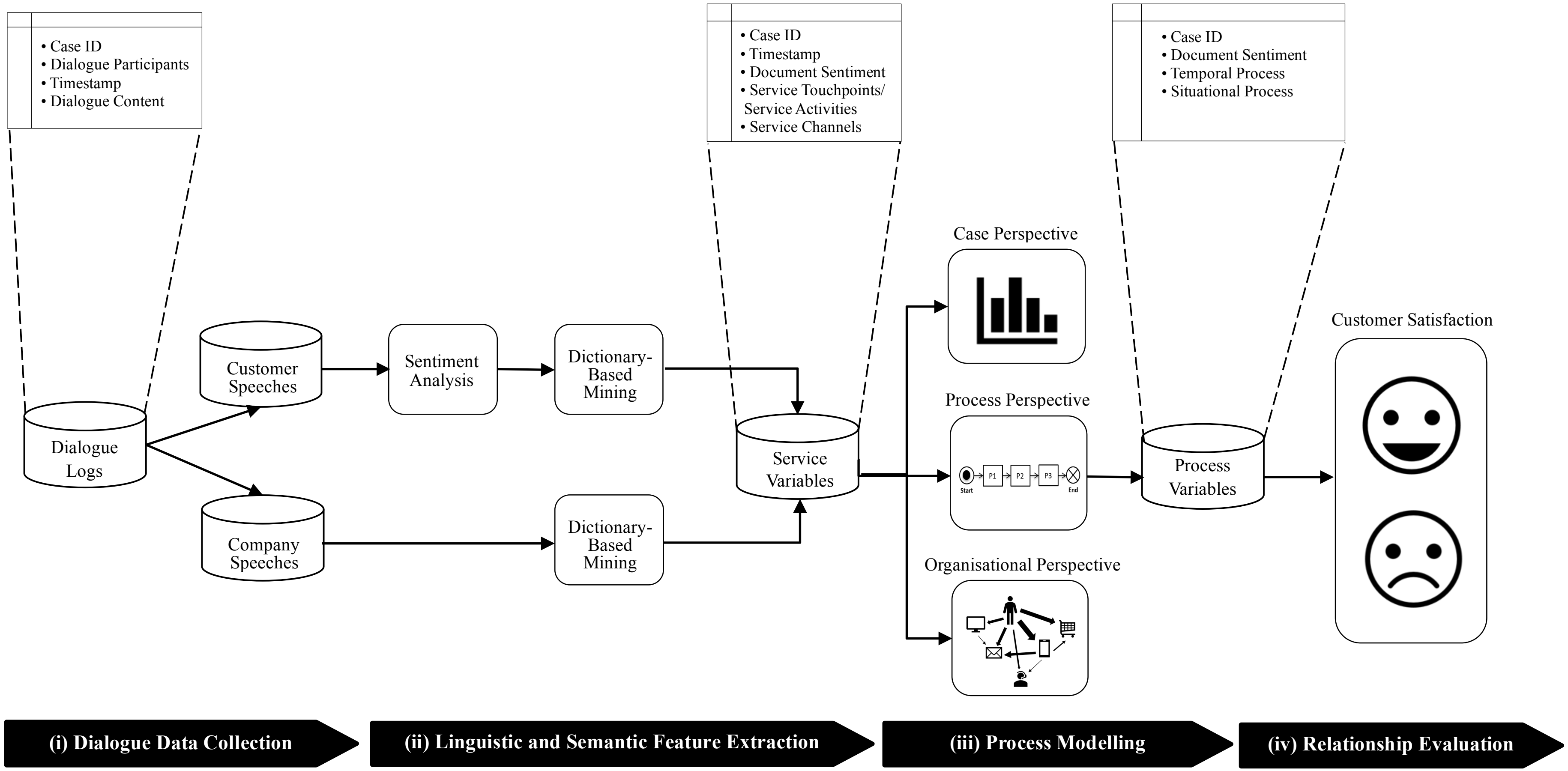

Figure 4.1 The Dialogue-Mining Apparatus

4.4.2 Linguistic and Semantic Feature Extraction

Analysis of the linguistic and semantic dimension of dialogues can be conducted using text mining (Ordenes et al. 2014). Text mining refers to the process of extracting insightful information and uncovering hidden knowledge within textual data (Ur-Rahman and Harding 2012). Several text-mining techniques have been introduced to classify user-generated content, such as linguistic approaches (Meijer et al. 2014; Ordenes et al. 2014), statistical approaches (Farhadloo et al. 2016) and hybrid approaches (Tirunillai and Tellis 2014). In this paper, a linguistic approach is adopted to analyse service recovery dialogues.

Linguistic approaches facilitate natural language processing (NLP) such as part-of-speech (POS) tagging and sentiment tagging to extract information in documents (Xu et al. 2011). Ordenes et al. (2014) stated that a collection of textual data is often related to particular domains (e.g. finance, biology), and to extract concepts from the dataset, domain-specific resources, such as lexicons and ontologies, can be used as the prior knowledge. Since service recovery is a mature research area with ample research findings regarding service failures, complaint-handling strategies, and customer post-recovery satisfaction (Buttle and Burton 2002; Fan and Niu 2016; Hart et al. 1990; Maxham 2001; Smith et al. 1999), linguistic-based extraction is considered a suitable method due to the well-established prior knowledge.

Two types of linguistic-based extraction were employed for examining service recovery dialogues; they are, sentiment analysis and dictionary-based mining. Sentiment analysis aims to classify textual documents, such as user reviews and social media, posts, into positive, negative, or neutral sentiments (Ordenes et al. 2017). The use of sentiment analysis allows for automated identification of customer complaints and non-complaints. Numerous sentiment lexicons have been developed to aid sentiment extraction, such as SentiWordNet (available from http://sentiwordnet.isti.cnr.it). The central idea of such lexicons is to assign each word a sentiment score and measure the sentiment strength by summing the total scores of a document (Singh et al. 2016).

Although sentiment analysis has become a popular method for identifying customer complaints, it provides limited understanding of critical factors during service interactions that trigger certain negative customer responses (Ordenes et al. 2014). Therefore, dictionary-based text mining was subsequently used to extract the information concerning product and service characteristics. Dictionary-based text mining relies on the development of text analysers or dictionaries that contain a set of pre-defined extraction rules. According to Franca and Sebastiani (2004), constructing a text analyser requires the prior knowledge $C = \{c_1, c_{2,\ldots\ldots} c_{|c|}\}$. In the dataset $\Omega = \{d_1, d_{2,\ldots\ldots} d_{|\Omega|}\}$, a training sample $Tr = \{d_1, d_{2,\ldots\ldots} d_{|Tr|}\}$ of the documents is taken from Ω to learn the characteristics of C. The results of learning are used to develop a text analyser and later applied to a testing sample $Te = \Omega - Tr$ to examine the degree of correspondence.

Similar to Ordenes et al. (2014), we developed text analysers that extracted information regarding service recovery using a small sample of dialogue data (*Tr)* to learn about the service experiences and service recovery interactions (C) in the whole dataset (Ω). This sample was manually annotated by human coders to fine-tune the text analysers before being applied to the remainder of the dataset (*Te*) for automated extraction.

4.4.3 Process Modelling

Process modelling aims to examine temporal processes and situational processes within service recovery dialogues (Finne and Grönroos 2009, 2017). The process dimension of dialogues can be uncovered using the process mining suggested by Schoor and Bannert (2012). Process mining assumes that it is possible to sequentially model event logs to re-construct business processes (van der Aalst 2011). In the information extraction phase, each complaint-handling dialogue was transformed into a structured format containing numerous variables related to specific products/services. The structured data can be viewed as event logs including information about: (i) a case (a complaint case), (ii) timestamps of actor speeches in a case, (iii) activities (e.g. service touchpoints), and (iv) resources (e.g. service channels) (van der Aalst et al. 2007; Rovani et al. 2015). Process mining can examine the event logs without using any a priori information to simulate a good characterisation of all possible paths (process discovery) or measure the alignment between the real process and the existing process models (process conformance) (van der Aalst 2011).

Three perspectives of process mining were highlighted by Song and van der Aalst (2008): case, process, and organisational. The case perspective evaluates the properties of the cases featured by their paths in the process, such as the frequency of a path (a set of activities) in a process model, and the average time between two activities (Song and van der Aalst 2008). The process perspective is helpful for investigating the temporal sequences of service processes. It aims to re-construct a sequenced process model using the timestamp of each business activity. Finally, the organisational perspective focuses on task performers involved in the event logs, exploring the situational processes, such as the relationships between task performers and the associations between tasks and task performers (Song and van der Aalst 2008).

4.4.4 Relationship Evaluation

Relationship is evaluated based on the experiences of actor interaction over time (Ballantyne and Varey 2006). Payne et al. (2008) stressed that the crucial constructs impacting the quality of relationships and if the relationship is continued include customer satisfaction, the degree of customer engagement and repeated purchase. Payne and Frow (2005) highlighted monitoring customer satisfaction and repurchase intention is one of the most important tasks to provides a holistic view of the relationship and assess business performance.

Analysing customers' positive and negative sentiments offers an initial understanding of customer satisfaction with service experience (Ordenes et al. 2014). During service recovery, the company and the complainers have several back-and-forth information exchanges, and successful customer recovery can be defined based on sentiment change, from negative to positive, within the dialogues (Fan and Niu 2016). A variable of customer satisfaction was constructed by analysing the sentiment change in a service recovery dialogue. If the final customer speech contains a positive sentiment, it is defined as demonstrating customer satisfaction with the dialogue outcomes (Fan and Niu 2016); if it does not contain a positive sentiment, it is regarded as demonstrating customer dissatisfaction with the dialogue outcomes.

In the following two sections, we demonstrate two applications for addressing issues with respect to process recovery and customer recovery utilising the dialogue-mining apparatus.

4.5 Application of the Dialogue-Mining Apparatus to Improve Process Recovery

The first application addressed issues related to process recovery and explored service touchpoints mentioned in the dialogues to identify customer experiences. This method was suggested by Ordenes et al. (2014), who analysed the service resources, activities, and context within customer feedback following a planned service process. Based on Ordenes et al. (2014), we took a further step to investigate the customer sentiment bonded with each service touchpoint as a means of distinguishing the weak links during service delivery processes.

4.5.1 Twitter Dialogue Data

The dataset used in this application was collected from the Twitter webcare of a UK grocery retailer over a six-month period. Twitter is considered a suitable research platform to study service recovery for two reasons: firstly, Twitter is widely used as a focal customer service platform, with more than 70% of Fortune 500 companies operating an active Twitter account (Ratliff and Kunz 2014); secondly, as a microblog service platform, each Twitter post (known as a tweet) only allows 140 characters. This unique communication style makes tweets more similar to real-life conversations than the dialogues on other social media (e.g. user comments on Facebook).

The initial dataset contained 7,201 dialogues. After examining the duration of dialogue interactions, we found that about 88% of dialogues are completed within 10 days. Dialogues of more than 10 days might contain multiple cases, as a customer may interact with Twitter agents several times for different purposes. Thus, we excluded dialogues exceeding 10 days, which left a total of 6,149 valid dialogues, including 14,328 tweets sent by customers and 10,810 tweets sent in response by the company.

To understand the embedded customer experience, we isolated customer tweets from dialogues and carried out text mining to extract customer sentiment and service touchpoints in the customer tweets (Step 2). Then, we mapped the service processes (Step 3) and compared positive and negative customer experiences (Step 4) to identify strengths and weaknesses of the company's service process.

4.5.2 Information Extraction from Customer Tweets

Sentiment Tagging on Customer Tweets

According to Verhoef et al. (2009), customer experience is a multidimensional construct including customers' cognitive, emotional, behavioural and social responses to companies' offerings. We used sentiment analysis to extract customer feelings embedded in their tweets to evaluate customer experience. Several publically shared sentiment lexical resources have been employed in prior research to mine user-generated content on social media (Ordenes et al. 2017). However, such sentiment lexicons were constructed to provide an out-of-context analysis (the sentiment of a word was assessed without considering the contexts) and were also limited to examining domain-specific data or controversial meanings in documents, such as sarcasm (Lau et al. 2014). Also, the criteria of sentiment polarity differ amongst lexicons.

To obtain objective results for customer tweet classification, three sentiment lexicons were used: SentiWordNet[5], SentiStrength[6] and the method proposed by Kolchyna et al. (2016), which adds emoticons into the sentiment lexicon for tweet

5 SentiWordNet is a lexical resource for opinion mining, and the current version is SentiWordNet 3.0 (available from http://sentiwordnet.isti.cnr.it)

6 SentiStrength is a sentiment analysis program in social web text, which supports 14 languages (available from http://sentistrength.wlv.ac.uk).

analysis. We used Python programming to perform the sentiment analysis. At the first stage, we removed the stop words (e.g. “the”, “and”) and conducted word stemming to return the words to the root form. Subsequently, a customer tweet was separated into words and assigned sentiment scores. We obtained the final sentiment score by calculating the total positive score minus the total negative score, and we assigned sentiment tags to each tweet (Singh et al. 2016). Notably, the three lexicons might result in different sentiment tags. We compared the resulting sentiment tags generated by these lexicons following this principle: only when two or more lexicons reached agreement on the sentiment (positive/negative/neutral) could the sentiment tag of a tweet be decided. For tweets with three different sentiment tags, human experts manually analysed the tweet content and gave appropriate tags. In the sentiment analysis phase, we identified 4,492 positive sentiment tweets, 6,396 neutral sentiment tweets, and 3,440 negative sentiment tweets.

Dictionary-Based Extraction for Customer Experience

Customer experience relates to a wide range of the company’s offerings, such as brand image, frontline employees, and platforms, and the overall experience comprises distinct touchpoints (Homburg et al. 2017). A collection of touchpoints forms customer experience journeys (Homburg et al. 2017; Lemon and Verhoef 2016). These journeys are dynamic, as customers can actively choose, participate, skip, or exit at any point planned by service providers (Heinonen et al. 2010). To understand customer experience embedded in tweets, we constructed a dictionary containing extraction rules regarding the key touchpoints. A two-phase iteration was adopted to develop the dictionary following Ordenes et al. (2014).

In the first iteration, we selected a random sample of 1,000 customer tweets as the training sample. Two annotators manually coded the words, phrases or sentences into concepts relevant to specific touchpoints to build up the initial dictionary. For example, we coded "charging" into *Paying* and "found a bug in my salad" into *Product/Service Using* (see Table 4.1). In this iteration, we captured approximately 300 concepts from the training sample. Also, we consulted the retailer's website, capturing another 1,000 concepts regarding its product range and services. These concepts were classified into 13 main touchpoints of customer experience: *Advertisement*, *Information Request*, *Product Sourcing*, *Product Stocking*, *Price*, *Package and Labelling*, *Deal and Promotion*, *Paying*, *Shopping Environment*, *Staff Support*, *Product/Service Using*, *(Non)Recommendation* and *After-Sales Contact*.

The initial dictionary was revised by lab labelling sessions – the second iteration. Six coders with marketing and management backgrounds were hired to analyse the customer tweets. Participants were incentivised and received a fixed payment, as recommended by Kuehl et al. (2016). We randomly selected 300 tweets from the training sample for collective labelling. In the sessions, the participants were paired into three groups and worked on the same 100 customer tweets following the coding template developed in the first iteration (the 13 touchpoints). The labelling sessions provided an objective judgement on the text-indexing rules. We compared the agreement between coders and corrected conflicting coding (73%, 71% and 60% in the coding groups, respectively). In this way, we improved the primary dictionary.

The manually crafted rules were then applied to build a text analyser in the software IBM SPSS Modeler 14.2. We used the software to conduct an automated classification of the entire set of customer tweets, with the exception of the 1,000 tweets used for training, and assigned the tweets to the 13 touchpoints. Moreover, the software facilitates the statistical approach – TF-IDF (term frequency–inverse document frequency) – to extract the important terms that were not found in the manual annotation. In this way, we continuously improved the text analyser by adding new terms. To evaluate the performance of automated information extraction, we calculated the data accuracy that represents the number of correct classifications based on the dictionary (Ordenes et al. 2014).

Table 4.1 Examples of Manual Annotation on the Training Sample

Tweet Content	Extracted Terms	Sub-Categories	Main Categories
Customer Tweets			
Do you have a company strategy of charging more at the till than shown on shelves? Overcharged again today!	charging, Overcharged	Payment	Purchasing Stage
Found a bug in my salad #yuck I've already emailed customer services with all the pics - but not heard back	Found a bug in my salad	Product/Service Using	Post-purchasing Stage
	emailed customer services	After-Sales Contact	Post-purchasing Stage
Company Reply Tweets			
If you're not happy with the product, please return them to your local store :-)	please return them	Refund and Replacement	Recovery Activity
	local store	Local Store	Service Channel
We would like to make this right for you. Can you call our Careline team on (number)?	call our Careline team	Channel Directing	Recovery Activity
	Careline team	Hotline	Service Channel

4.5.3 Mapping Customer Experience Processes

In prior research (e.g. Homburg et al. 2017; Lemon and Verhoef 2016), customer experience journeys were often divided into three major stages: pre-purchase (e.g. brand recognition, information searching), purchase (e.g. product choice, ordering, paying) and post-purchase (e.g. product consuming, word of mouth, future purchase). We used the three stages of the customer journey (Lemon and Verhoef 2016) as a priori information and mapped customer tweets containing the touchpoints into pre-purchase, purchase, and post-purchase.

4.5.4 Comparison between Positive and Negative Customer Experiences

Identifying key touchpoints in customer journeys and the customer feelings bonded with the touchpoints helps reveal customer satisfaction or dissatisfaction with specific service processes. We divided customer tweets into positive and negative customer experience sets based on the sentiment tags and compared the frequency of each touchpoint mentioned in these two types of customer experience journeys. Through such comparison, we evaluated the relationship dimension of dialogues and offered insights into the company's strengths and weaknesses during service delivery processes.

4.5.5 Results

Table 4.2 shows the results of the automatic information extraction from the customer tweets. As shown, the accuracy is between 94% and 50%, with 10 out of 13 categories higher than 70%. Notably, each tweet may contain more than one touchpoint, and thus the sum of the tweet count per category (7,830) is more than the total number of customer tweets (6,149).

Of the captured categories of customer experience, the text analyser identifies 2,432 tweets referring to the touchpoints in the purchase phase, 2,169 tweets regarding the post-purchase phase and only 332 tweets regarding the pre-purchase phase. *Product/Service Using* is found to possess the highest number of tweets (2,174), followed by *Shopping Environment* (810) and *Staff Support* (622). However, there are 1,315 tweets for which our text analyser failed to identify any concepts related to the touchpoints. The uncategorised tweets often contain only customers' emotional responses, such as "I love you, @company", weblinks or pictures that explain the details of their service experience, such as "This is not right at all, (weblinks)", or non-transactional activities (e.g. charity donations). These tweets cannot be classified by the text analyser.

Table 4.2 Resulting Touchpoints of Customer Experience

Customer Experience Touchpoints	Count	Accuracy (%)
Categorised Tweets		
Pre-purchase	*332*	
Advertisement	273	90.74
Information Request	88	51.61
Purchase	*2,432*	
Product Sourcing	98	83.33
Product Stocking	490	75.58
Package and Labelling	160	50.00
Deal and Promotion	532	75.00
Price	269	81.48
Shopping Environment	810	78.81
Staff Support	622	69.47
Paying	449	88.89
Post-purchase	*2,169*	
Product/Service Using	2,174	93.50
(Non)Recommendation	49	90.90
After-Sales Contact	501	79.16
Uncategorised Tweets	1,315	-
Total	6,149	-

The results of the process analysis offer an overview of the customer journeys in the context of positive and negative service experiences. We compared the number of positive tweets and negative tweets at each touchpoint and verified the results using chi-square analysis. As shown in Figure 4.2, customers tended to have negative experiences at the touchpoints of *Product/Service Using*, *Paying*, *Staff Support*, *Shopping Environment*, *Price, Package and Labelling* and *Product Stocking* ($p < 0.001$). The highest percentage of negative experience is found in *Product/Service Using*, where customers frequently mentioned the concepts "out of date" relating to product expiration and "awful taste" regarding product quality. In contrast, positive customer experiences are shown in *Advertisement*, *Deal and*

Promotion, and *(Non)Recommendation* (p <0.001). In exploring the positive experiences, we found that numerous customers shared positive word-of-mouth the company's seasonal advertisement by tweeting "the best Christmas ad" or engaged in a hashtag event (#tasteandtell) of product recommendation.

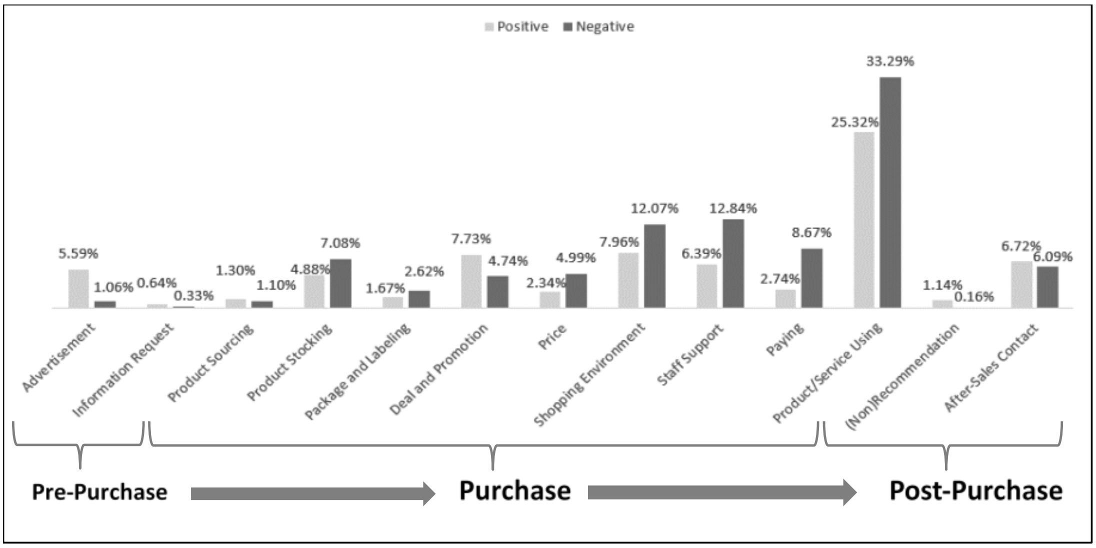

Figure 4.2 Comparison between Positive and Negative Customer Experiences

4.6 Application of the Dialogue-Mining Apparatus to Improve Customer Recovery

The second application aims to address the issue of customer recovery by examining the tweets of both the company and the customers. The results of this application can help practitioners to identify effective and ineffective activities and processes for recovering customer satisfaction.

4.6.1 Twitter Dialogue Data

According to Fan and Niu (2016), each piece of Twitter dialogue can be treated as a "mini case" involving conversation exchanges between a customer and a Twitter service agent. A complaint-handling dialogue is completed when the outcome is observable, with either the complaint being addressed (successful recovery) or the complainant exiting without resolution (failed recovery). In the first application, we assigned sentiment tags to each customer tweet to classify positive and negative customer experiences. We selected the cases containing negative sentiments as customer complaint cases. Specifically, we focused on the issue occurring in the touchpoint *Product/Service Using* for further analysis. A total of 650 service recovery dialogues were identified.

Again, the proposed dialogue-mining apparatus was applied to analyse the activities of service recovery mentioned in the company's tweets (Step 2). The embedded activities were later mapped into the customer recovery processes (Step 3). Finally, we compared between the activities and processes within customer satisfaction and dissatisfaction dialogues to identify the poor practices of service recovery (Step 4).

4.6.2 Information Extraction from Service Recovery Dialogues

Sentiment Tagging on Customer (Dis)Satisfaction Dialogues

In the service recovery dialogue dataset, each customer tweet was assigned a sentiment tag in the first application. To understand customer post-recovery satisfaction, we analysed the sentiment change in dialogues (Chebat and Slusarczyk 2005; Fan and Niu 2016). If the sentiment of a service recovery dialogue changed from a negative sentiment in the first customer tweet to a positive sentiment in the final customer tweet, it was considered a customer satisfaction case (Fan and Niu 2016), but if the final customer tweet remained negative or showed a neutral sentiment, it represented customer dissatisfaction with the service recovery. Based on this principle, we identified 124 customer satisfaction dialogues.

Dictionary-Based Extraction for Service Recovery Activities and Channels

We repeated the two-phase iteration in the previous application. Two domain experts manually coded the training sample of 1,000 company tweets to identify the key service recovery activities and service channels mentioned in the company tweets. Table 4.1 provides another example of the coding scheme based on the company's reply tweets. As shown, we coded "please return them" into *Refund and Replacement* and "call our Careline" into *Channel Directing*. Moreover, we also coded the terms representing service channels, such as "Careline" coded into *Hotline*. Similarly, the initial dictionary of service recovery was revised by lab labelling sessions to achieve objectivity. As a result, approximately 600 concepts were captured and assigned to 11 types of service recovery activities and eight types of service channels.

Based on the manual annotation results, we built another text analyser of service recovery activities and service channels in SPSS Modeler. We performed the automated information extraction and classification on the company reply tweets and evaluated the accuracy.

4.6.3 Mining Customer Recovery Processes

In the process-mining phase, we modelled both the temporal process and situational process of customer recovery. The temporal process explored the sequences of service recovery activities to root out inefficient processes that may lead to customer exit and other negative outcomes on social media. The situational process considered the complaint channel management by revealing the associations between service recovery activities and service channels. Such process analysis allows the company to understand the position of each channel in complaint handling.

Temporal Process

A process map that investigates and visualises the sequences of the observed activities embedded in the event logs is useful for examining the temporal process (Sedrakyan et al. 2016). To unearth customer recovery processes, we employed the process discovery software Disco[7]. The software, facilitating the case perspective and process perspective, can construct process maps without using any a priori information, and it also evaluates the time interval between pairs of activities in the process. This allows us to assess the efficiency of the delivery of service recovery activities.

The event logs used in process mining were created in Step 2, in which we extracted service recovery activities in the company tweets. Table 4.3 provides two examples of Twitter event logs. As shown, case A111 represents a customer satisfaction case regarding the issue of *Poor Product Quality*. A111 contains three dialogue interactions, within which three types of service recovery activities are present. By sequentially arranging the timestamps in A111, we obtained a service recovery path of *Refund and Replacement − Channel Directing – Guarantee*. Using the process-map discovery tool, the models manifested themselves based on the timestamps and the corresponding activities (Sedrakyan et al. 2016). By aggregating all paths in the Twitter event logs, a holistic view of the service recovery process can be unearthed.

7 Disco is a commercial tool for process mining developed by Fluxicon: http://fluxicon.com/disco/.

Situational Processes

We further applied the organisational perspective to analyse the situational process where task performers were involved in event logs (Song and van der Aalst 2008). In Step 2, we extracted service channels (task performers) mentioned by the Twitter agents. As shown in Table 4.3, case A333 contains three channels: *Twitter*, *Local Store,* and *Website*. Each service recovery channel has a corresponding service activity. In other words, the initial service recovery activity was conducted by Twitter, and then Twitter agents would engage other channels in certain recovery activities to resolve customer complaints together. To explore such channel-activity relationships, we applied correspondence analysis that analyses two-way contingency tables of categorical data to visualise the interrelationships of variables on a two-dimensional map (Carroll et al. 1986). We assigned service recovery activities into rows and service channels into columns. Since there is a correspondence between the row and column coordinates, both variables can be plotted onto the same joint space and interpreted by visually examining their positions and closeness (Calantone et al. 1989).

Table 4.3 Examples of Twitter Dialogue Variables

Case ID	Timestamp	Service Failure	Service Recovery Activity	Service Recovery Channel	Service Recovery Outcome
A111	26/06/15 10:54:16	Poor Quality Product	Refund and Replacement	Local Store	Satisfaction
A111	26/06/15 13:35:24	Poor Quality Product	Channel Directing	Hotline	Satisfaction
A111	27/06/15 11:20:35	Poor Quality Product	Guarantee	Twitter	Satisfaction
A333	21/03/15 14:05:05	Staff Mistreatment	Apology	Twitter	Dissatisfaction
A333	21/03/15 14:05:05	Staff Mistreatment	Channel Directing	Local Store	Dissatisfaction
A333	26/03/15 12:55:41	Staff Mistreatment	Information/Explanation	Website	Dissatisfaction

4.6.4 Examining Customer Satisfaction with Service Recovery

We examined customer satisfaction with service recovery at the levels of service recovery activities and processes. Firstly, we compared the frequency of each service recovery activity in both the customer satisfaction and dissatisfaction dialogue subsets. Then, to understand if a sequence in which the company implemented certain service recovery activities in dialogues can be indicative of customer post-recovery satisfaction, we compared the process models of these two dialogue subsets. Such comparisons reveal the difference between these two service recovery outcomes and offer insights into service recovery operations.

4.6.5 Results

Table 4.4 provides the results of information extraction regarding service recovery activities and service channels in the context of the *Product/Service Using* issue. The text analyser successfully classifies 573 cases (88.15%), with most of the resulting categories achieving accuracy higher than 70%, with the exception of the categories *Information / Explanation* (56.52%) and *Website* (53.33%). In more than half of the dialogues, the agents directly handled service problems on Twitter (315 cases) by asking the complainants to provide pictures of product barcodes or receipts. Only 61 cases contain the *Channel Directing* activity through which the company migrated complainants to other channels. Other high-frequency activities mentioned by the Twitter agents include *Showing Empathy* (299 cases) and *Process Control* (219 cases). These findings from the information extraction hold true in both the customer satisfaction and dissatisfaction subsets. The chi-square analysis indicates no significant differences between successful and non-successful recovery throughout each service recovery activity and service channel (an exception is the activity *Showing Empathy*, with p =.028).

Table 4.4 Extracted Categories of Recovery Activities and Channels

	Count	Customer Satisfaction	Customer Dissatisfaction	Accuracy (%)
Categorised Cases	**573**	**108**	**455**	-
Service Recovery Activities				
Refund and Replacement	144	34	106	94.44
Compensation	34	12	19	88.23
Problem Handling	315	89	303	99.36
Process Control	219	63	214	90.87
Channel Directing	61	11	36	90.16
Follow-Up	71	16	53	84.50
Information / Explanation	46	7	43	56.52
Guarantee	49	10	31	93.87
Active Help	8	0	7	100
Apology	137	35	116	100
Showing Empathy	299	57	257	97.32
Service Channels				
Local Store	151	46	136	82.78
Customer Care Team	79	23	74	96.20
Supplier	79	24	67	100
Specific Department	62	20	42	77.41
Franchise	24	6	18	83.33
Website	15	3	7	53.33
Hotline	7	0	4	71.43
Email	1	0	0	100
Uncategorised Cases	77	16	71	-
Total	650	124	526	-

In the process analysis, a comparison drawing on the process models was conducted to examine the difference between customer satisfaction cases and customer dissatisfaction cases. The uncategorised cases were excluded when conducting process modelling as these cases contain tweets with no relevant service recovery activities and service channels mentioned. The satisfaction subset contains 108 cases with 334 events spread out over 11 service recovery activities, while the dissatisfaction subset contains 455 cases with 1,185 events. We applied multiple filtering rules to obtain process maps. The process-mining software enables the processes to be defined at different levels of abstraction by removing the low-frequency activities and paths, which helps solve the "spaghetti problem" and improves the power of interpretation (van der Aalst and Günther 2007). We tested different filtering rules and constructed the final models by setting the preserve threshold as 60% of the most frequent activities and 20% of the most frequent paths to retrieve simplified (but not over-simplified) service recovery processes. Based on this condition, we compared the process models of satisfaction and dissatisfaction cases, and the resulting models are in Figure 4.3 and Figure 4.4, respectively.

As shown, the numbers in the boxes represent the frequency of service recovery activities, and the numbers on arrows refer to the frequency of the connection between two sequenced activities and the time intervals between them. As the models were built based on higher abstraction levels, the numbers can only reflect parts of the paths. These two process models share some similar patterns. *Showing Empathy* is identified as the initial activity that the company delivered to complainants, and the most important subsequent activity is *Problem Handling*. Similarly, another major path found in both the satisfaction and dissatisfaction models is *Showing Empathy – Refund and Replacement – Problem Handling*.

In terms of the difference, we specifically highlight three observations from the models. Firstly, the satisfaction subset shows more efficiency in the *Problem Handling* path – that is, following *Problem Handling,* the company either enters a loop of *Compensation – Problem Handling* or implements a *Follow-Up* activity before completing the process. In contrast, in the dissatisfaction subset, after *Problem Handling*, the recovery process enters a long path with several loops involved. The loops represent the same activity conducted more than once, which indicates the reduced efficiency in service recovery.

Secondly, both models contain a non-overlapped activity. As shown, the activities in both models are mostly the same, but in the satisfaction recovery model, *Compensation* is present in the later phase of the process. *Compensation* is the actual offering (e.g. vouchers) provided by the company. Also, in the dissatisfaction model, we find the activity *Channel Directing* shown in the early path of the process, yet it is absent in the satisfaction model. Perhaps not surprisingly, the paths involving *Channel Directing* tend to be longer than other paths.

Thirdly, the time intervals between two activities are found to be longer in the dissatisfaction model than the satisfaction model. Our models provide the details of the time interval between the sequenced pairs of activities using median duration, as the dataset is not normally distributed. In the satisfaction recovery model, the longest time interval is 24.3 hours present in a self-loop of *Process Control* activity. In contrast, the longest time interval in the dissatisfaction model is 67.1 hours, which is shown in the path *Follow-Up – Showing Empathy*.

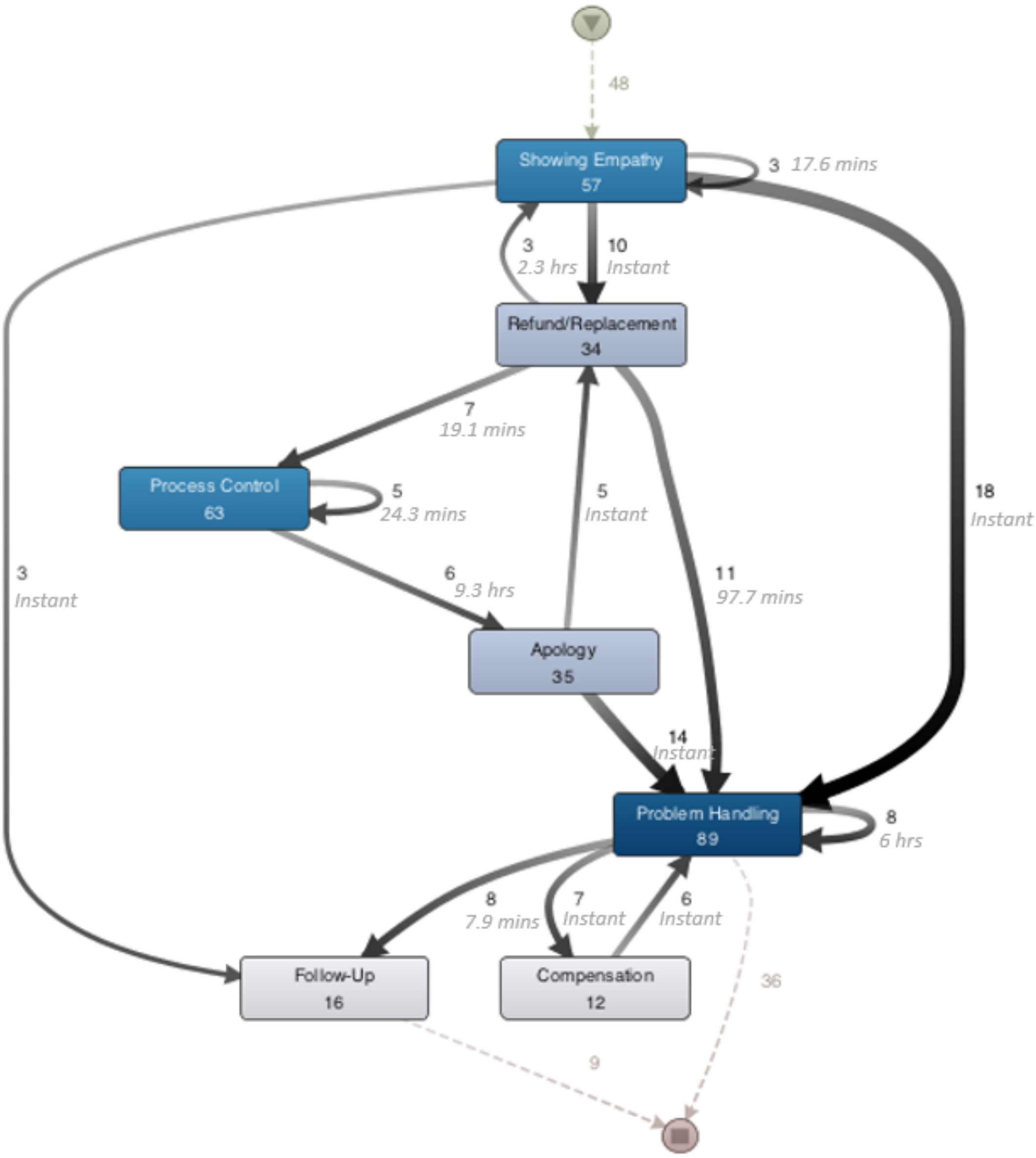

Figure 4.3 The Service Recovery Process of Customer Satisfaction Cases

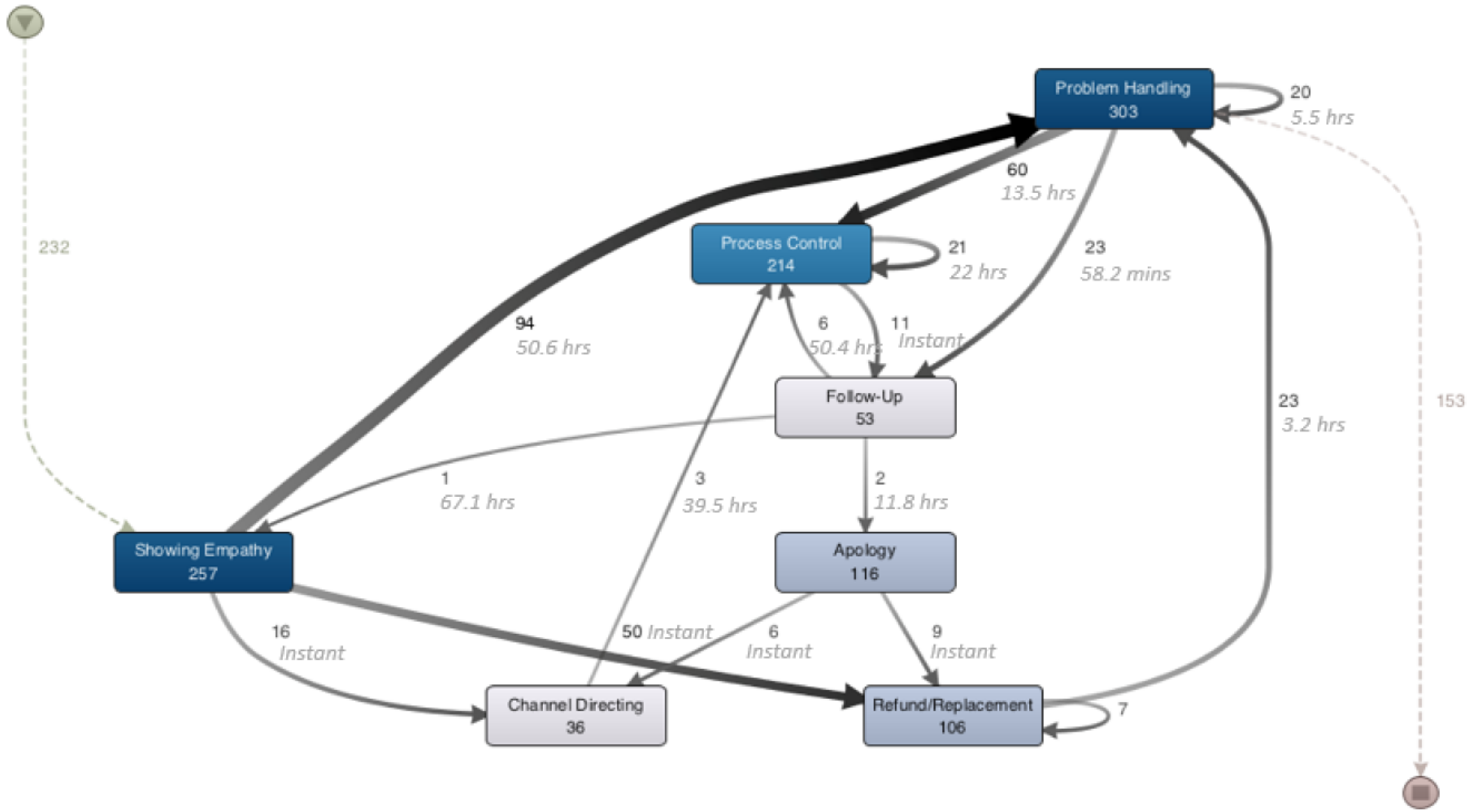

Figure 4.4 The Service Recovery Process of Customer Dissatisfaction Cases

When it comes to the service recovery channels, Table 4.4 shows that *Local Store* possesses the highest engagement (151 cases) during service recovery, followed by *Supplies* (79) and *Customer Care Team* (79). In contrast, *Hotline* (7) and *Email* (1) are the least mentioned channels during service recovery. We further examined the associations between service recovery activities and service channels using correspondence analysis. The result shows a significant dependency between these two variables (chi-square [70df] = 2617.253, $p < .001$). The correspondence analysis portrays the 11 types of service activities (rows) and nine types (eight identified channels and Twitter) of service channels (columns) in a joint space, thus revealing the relationships between service activities and channels. Figure 4.5 demonstrates the graphical output of the correspondence analysis result in the context of *Product/Service Using* issue. An apparent trend is present in the upper right side, with three data points – *Hotline*, *Franchise* and *Channel Directing* – widely separated from the rest of the points. This indicates that Twitter agents tended to migrate complainants (*Channel Directing*) to the channels *Hotline* and *Franchise*. The second trend is the closeness of *Local Store* and *Refund and Replacement*. These two are slightly distant from the majority of data points. The results suggest the venue where supportive recovery activities occur, enabling companies to plan resource allocation among service recovery channels.

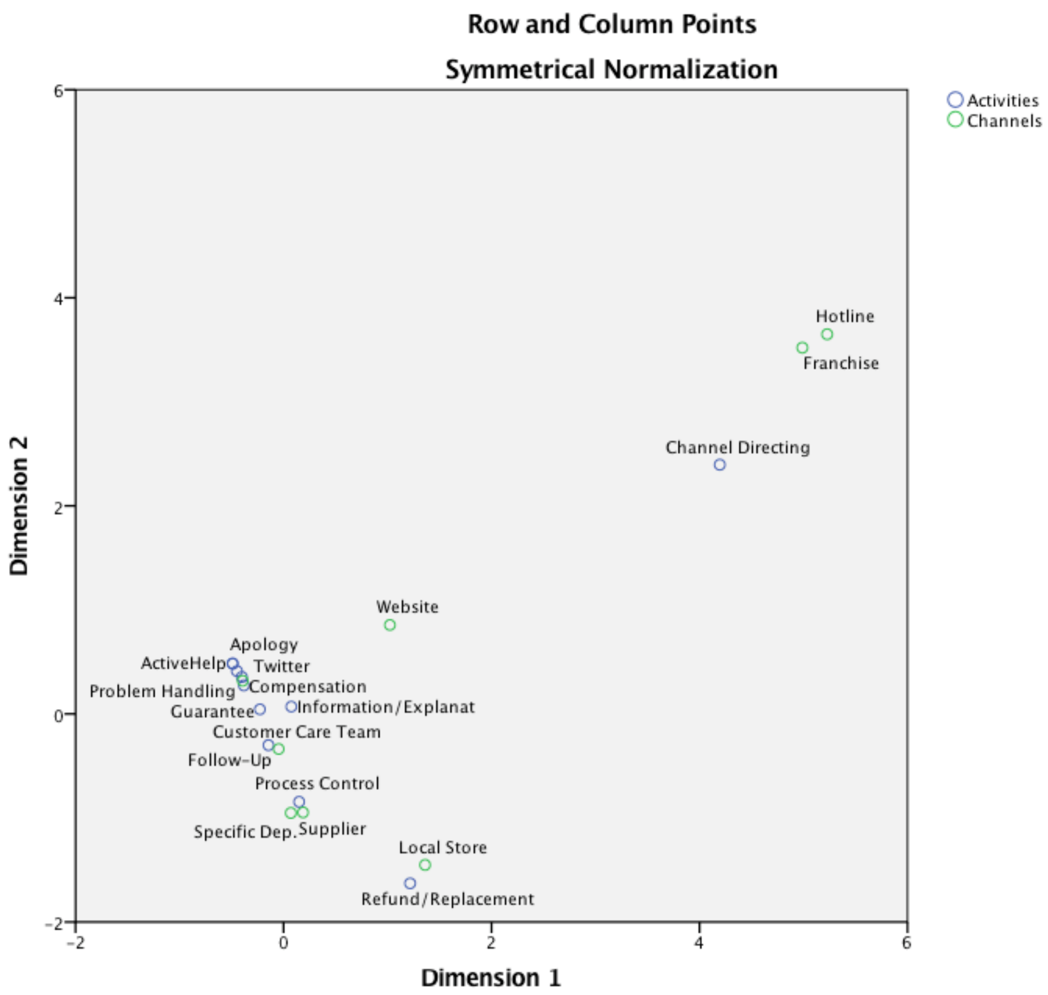

Figure 4.5 Correspondence Analysis of Service Recovery Activities and Service Channels

4.7 Discussion

4.7.1 Theoretical Implications

The proposed framework is established on the value co-creation literature, evaluating the linguistic and semantic dimension, process dimension, and relationship dimension embedded in the interactive dialogues. Ordenes et al. (2014) suggested that such an analytical framework should be viewed as a theorised model that provides opportunities of open learning and can be continuously advanced by subsequent research. The linguistic and semantic dimension is designed to explore and evaluate the "acts" and "episodes" within service dialogues (Grönroos 2004). In the first application, we defined the “acts” as the service touchpoints, which relate to the products, services, platforms, or employees encountered by customers (Homburg et al. 2017). Yet, in the second application, the touchpoint *Product/Service Using* was viewed as an “episode”, involving many “acts” of service recovery activities and service channels.

The aggregated features at the linguistic and semantic level form a process flow. Ordenes et al. (2014) showed an attempt to map the service features embedded in customer feedback into a service process. Their work sheds light on understanding customer experience and value co-creation following a pre-defined process. In the same vein as Ordenes et al. (2014), the first application models customer tweets following the pre-defined stages of customer journeys (Lemon and Verhoef 2016), and it uncovers the strengths and weaknesses in the company’s service delivery process. Such method is useful to understand customer value-creating process, where customers use activities and resources to manage their business and relationships with companies (Payne et al. 2008). However, service interaction processes in the real-world setting tend to be more dynamic as customers (and companies) can actively choose to participate and exit at any stage of the service process (Heinonen et al. 2010). In addition to investigating the pre-defined processes, our work provides an advanced technique to analyse dynamic service processes. In the second application, we modelled both temporal processes and situational processes without using prior knowledge. The dynamic process models enable companies to understand interaction-driven service processes and offer insights into dynamic service process management. Payne et al. (2008) called such dynamic process as encounter processes, where the interaction and exchange could be better managed to develop successful co-creation opportunities.

In the relationship dimension of dialogues, we examined customer satisfaction based on sentiment change during dialogue interactions (Fan and Niu 2016). According to Grönroos (2004), the customer relationship is developed through "sequences" of "acts" and "episodes". More specifically, the customer relationship should be viewed as the outcome of dialogues. Examining the impacts of the lower-level factors, such as service recovery activities (linguistic and semantic dimension) and processes (process dimension), on the relationship is suggested as an important approach to manage dialogue interactions. This is because customers' previous interactions with the company tend to have impacts on the next stage of interactions (Lemon and Verhoef 2016) and ultimately influence the long-term customer relationship.

4.7.2 Managerial Implications

The dialogue-mining approach serves as an improved solution for examining service recovery dialogues and allows for efficient identification of poor practices during service recovery via social media. In the first application, we demonstrated an advanced dialogue-mining method to perform process recovery by analysing customer complaints. Prior research suggested that data for process recovery can be collected from customer surveys in which customers disclose the critical incidents of service failures (Johnston and Michel 2008). Customer feedback on social media offers good-quality critical incidents without the drawbacks of qualitative data collection, such as participants' memory errors and catering to researchers' expectations. Importantly, this data is present in a real-time format, even while the service issue is still happening (Abney et al. 2017). In this way, the dialogue-mining method can support practitioners in understanding customer experiences speedily and making decisions on process recovery in an agile manner.

The second application sheds light on managing service recovery activities and processes to improve the practice of customer recovery via social media webcare. Managing interactive dialogues becomes an important business strategy as customer recovery via social media often takes place in the public arena where active and passive participants are present, and the recovery outcomes can influence not only the complainers but also their associates in the network (Schaefers and Schamari 2016). Marketing practitioners have shown interest in studying speeches of frontline employees to avoid negative customer outcomes (Dixon et al. 2010). Our work highlights that, in addition to the content relating to customer recovery activities, the processes that the service recovery activities delivered to complainers, such as sequences and channel migration, are influential for customer post-recovery satisfaction. The dialogue-mining method can help practitioners to evaluate consistency between the service agent’s actions and the company’s strategy, while also improving customer support by rooting out duplicated service processes, reducing response times, and improving resource allocation across different service channels.

Finally, the dialogue-mining method posits an analytical framework of dialogue collection, information extraction, process modelling, and relationship analysis. This method addresses the drawbacks of qualitative research methods in terms of time-consuming data analysis and small sample sizes being examined (e.g. Fan and Niu 2016). Text mining enables speedy information extraction from a vast volume of textual data based on fine-tuned text analysers (Ordenes et al. 2014). Although the text analysers are built by a qualitative approach in this research, it can be re-used, enhanced, shared and applied to new datasets. Therefore, the text analysers can maintain up-to-date extraction rules and provide an ongoing contribution to the focal research domain.

4.8 Limitations and Further Research

This research has several limitations and offers opportunities for future research. First, the Twitter dialogues being examined are only conducted in the context of one-to-one interactions between a company and a customer. The proposed approach is thus more suitable for analysing the dialogue data collected from Twitter webcare, call centre and online chat service. However, interactions between a company and multiple customers are common. For instance, a dialogue on Facebook may be a post containing multiple customers' comments. Future research should understand such one-to-many or even many-to-many dialogs and explore the relevant issues, such as a compan's specific service recovery responses in managing multi-actor dialogues.

Second, we note some obstacles in processing Twitter dialogues when using the dialogue-mining apparatus. In the applications, there are a high proportion of tweets (e.g. 21.38 % in the first application) unable to be successfully identified any concepts and categorised into service touchpoints using dictionary-based extraction. This is because the unique platform policy only allows Twitter users to post a short message (140 words). Therefore, a Twitter user's speeches in the same context might be divided into several short messages that only contain a weblink, or emoticons. We suggest that further research analysing Twitter dialogues can test different units of analysis, such as a single tweet, a number of tweets belonging to an interaction, or all tweets from the same customer, to get better results of information extraction.

Third, the dialogue-mining apparatus serves as an improvement of existing methods and models that address practical service recovery issues. The proposed approach is useful to understand customer self-disclosed information such as customer satisfaction, or customer positive word-of-mouth. Yet, customer relationship is a construct consisting of multiple aspects, such as customer acquisition, customer retention, customer word-of-mouth intention, and customer loyalty (Payne and Frow 2005). To obtain more detailed information of customer relationship, we suggest that dialogue data should be combined with other types of data sources, such as customer purchase history data and customer churn data. The external data can offer more accurate prediction of customer relationship and thus a deeper understanding of customer lifetime value can be obtained (Ordenes et al. 2014).

Finally, although the proposed framework manifests its value for offering a problem-solving solution to service recovery, the applications were only validated by a dataset from a single company in a single industry. Hence, the findings are specific to the research context. However, the dialogue-mining framework/apparatus is generalisable and applicable for future research, not restricted to study service recovery in other service sectors but to explore other dialogue contexts, such as marketing campaign, and CSR appeal.

4.9 References

van der Aalst, W.M.P., Reijers, H.A., Weijters, A.J.M.M., van Dongen, B.F., Alves de Medeiros, A.K., Song, M. and Verbeek, H.M.W. (2007), "Business Process Mining: An Industrial Application," *Information Systems*, 32(5), 713–732.

van der Aalst, W.M.P. (2011), *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, Berlin: Springer-Verlag

van der Aalst, W.M.P. and Günther, C.W. (2007), "Finding Structure in Unstructured Processes: The Case for Process Mining," In *Proceedings 7th International Conference on Application of Concurrency to System Design, ACSD 2007*, 3–12.

Abney, A.K., Pelletier, M.J., Ford, T.S. and Horky, A.B. (2017), "#IHateYourBrand: Adaptive Service Recovery Strategies on Twitter," *Journal of Services Marketing*, 31(3), 281–294.

Ballantyne, D. (2004), "Dialogue and its Role in the Development of Relationship Specific Knowledge," *Journal of Business & Industrial Marketing*, 19(2), 114–123.

Ballantyne, D. and Varey, R.J. (2006), "Introducing A Dialogical Orientation to the Service Dominant Logic of Marketing," In R.F. Lusch, and S. L. Vargo, eds, *The Service-dominant Logic of Marketing: Dialog, Debate, and Directions*, Armonk, NY: ME Sharpe, 224–235.

Bowen, D.E. and Johnston, R. (1999), "Internal Service Recovery: Developing A New Construct," *International Journal of Service Industry Management*, 10(2), 118–131.

Buttle, F. and Burton, J. (2002), "Does Service Failure Influence Customer Loyalty?" *Journal of Consumer Behaviour*, 1(3), 217–227.

Calantone, R.J., Di Benedetto, C.A., Hakam, A. and Bojanic, D.C. (1989), "Multiple Multinational Tourism Positioning Using Correspondence Analysis," *Journal of Travel Research*, 28(2), 25–32.

Carroll, J.D., Green, P.E. and Schaffer, C.M. (1986), "Interpoint Distance Comparisons in Correspondence Analysis," *Journal of Marketing Research*, 23, 271–280.

Chebat, J.-C. and Slusarczyk, W. (2005), "How Emotions Eediate the Effects of Perceived Justice on Loyalty in Service Recovery Situations: An Empirical Study," *Journal of Business Research*, 58(5), 664–673.

Dalla Pozza, I. (2014), "Multichannel Management Gets "Social"," *European Journal of Marketing*, 48(7/8), 1274–1295.

Davidow, M. (2003), "Have you Heard the Word? The Effect of Word of Mouth on Perceived Justice, Satisfaction and Repurchase Intentions Following Complaint Handling," *Journal of Consumer Satisfaction, Dissatisfaction & Complaining Behavior*, 16, 66–68.

Dixon, M., Freeman, K. and Toman, N. (2010), "Stop Trying to Delight Your Customersl," *Harvard Business Review*, 88(7-8), 116– 22.

Einwiller, S.A. and Steilen, S. (2014), "Handling Complaints on Social Network Sites – An Analysis of Complaints and Complaint Responses on Facebook and Twitter Pages of Large US Companies," *Public Relations Review*, 41(2), 195–204.

Fan, Y. and Niu, R.H. (2016), "To Tweet or not to Tweet? Exploring the Effectiveness of Service Recovery Strategies Using Social Media," *International Journal of Operations & Production Management*, 36(9), 1014 – 1036.

Farhadloo, M., Patterson, R.A. and Rolland, E. (2016), "Modeling Customer Satisfaction from Unstructured Data Using A Bayesian Approach," *Decision Support Systems*, 90, 1–11.

Finne, Å. and Grönroos, C. (2009), "Rethinking Marketing Communication: From Integrated Marketing Communication to Relationship Communication," *Journal of Marketing Communications*, 15(2-3), 179–195.

Finne, Å. and Grönroos, C. (2017), "Communication-in-Use: Customer-Integrated Marketing Communication," *European Journal of Marketing*, 51(3), 445–463.

Fornell, C. and Wernerfelt, B. (1987), "Defensive Marketing Strategy by Customer Complaint Management: A Theoretical Analysis," *Journal of Marketing Research*, 24(4), 337–346.

Franca, D. and Sebastiani, F. (2004), "Supervised Term Weighting for Automated Text Categorization," In S. Sirmakessis, eds, *Text Mining and its Applications*: *Results of the NEMIS Launch Conference.* Berlin: Springer-Verlag, 81–97.

Gentile, C., Spiller, N. and Noci, G. (2007), "How to Sustain the Customer Experience: An Overview of Experience Components that Co-create Value With the Customer," *European Management Journal*, 25(5), 395–410.

Grainer, M., Noble, C.H., Bitner, M.J. and Broetzmann, S.M. (2014), "What Unhappy Customers Want", *MITSloan Management Review*, 55(3), 31–35.

Grégoire, Y., Salle, A. and Tripp, T.M. (2014), "Managing Social Media Crises with Your Customers: The Good, the Bad, and the Ugly", *Business Horizons*, 58(2), 173–182.

Grewal, D., Roggeveen, A.L. and Tsiros, M. (2008), "The Effect of Compensation on Repurchase Intentions in Service Recovery," *Journal of Retailing*, 84(4), 424–434.

Grönroos, C. (1988), "Service Quality: The Six Criteria of Good Perceived Service Quality," *Review of Business*, 9(3), 10–13.

Grönroos, C. (2004), "The Relationship Marketing Process: Communication, Interaction, Dialogue, Value," *Journal of Business & Industrial Marketing*, 19(2), 99–113.

Gummesson, E. and Mele, C. (2010), "Marketing as Value Co-creation through Network Interaction and Resource Integration," *Journal of Business Market Management*, 4(4), 181–198.

Hart, C.W., Heskett, J.L. and Sasser, W.E. (1990), "The Profitable Art of Service Recovery," *Harvard Business Review*, 68(4), 148–56.

Heinonen, K., Strandvik, T., Mickelsson, K., Edvardsson, B., Sundström, E. and Andersson, P. (2010), "A Customer-Dominant Logic of Service," *Journal of Service Management*, 21(4), 531–548.

Hirschmann, A.O. (1970), *Exit, voice and loyalty: Responses to Decline in Firms, Organizations, and State*, MA: Harvard University Press.

Holmlund, M. (1997), *Perceived Quality in Business Relationships,* Hanken Swedish School of Economics Finland/CERS, Helsingfors.

Homburg, C., Jozić, D. and Kuehnl, C. (2017), "Customer Experience Management: Toward Implementing an Evolving Marketing Concept," *Journal of the Academy of Marketing Science*, 45(3), 377–401.

Hwong, Y.L., Oliver, C., van Kranendonk, M., Sammut, C. and Seroussi, Y. (2017), "What Makes You Tick? The Psychology of Social Media Engagement in Space Science Communication," *Computers in Human Behavior*, 68, 480–492.

Johnston, R. and Michel, S. (2008), “Three Outcomes of Service Recovery - Customer Recovery, Process Recovery and Employee Recovery,” *International Journal of Operations & Production Management*, 28(1), 79–99.

Kolchyna, O., Souza, T. T. P., Treleaven, P. C. and Aste, T. (2016), “Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination,” In M. Gautam and Y. Xiang, eds, *Handbook of Sentiment Analysis in Finance*. Albury Books.

Kuehl, N., Scheurenbrand, J. and Satzger, G. (2016), “NEEDMINING: Identifying Micro Blog Data Containing Customer Needs,” In *Twenty-Fourth European Conference on Information Systems (ECIS2016)*. İstanbul, Turkey.

La, K.V. and Kandampully, J. (2004), “Market Oriented Learning and Customer Value Enhancement through Service Recovery Management,” *Managing Service Quality*, 14(5), 390–401.

Larivière, B., Joosten, H., Malthouse, E.C., van Birgelen, M., Aksoy, P., Kunz, W.H. and Huang, M.-H. (2013), “Value Fusion: The Blending of Consumer and Firm Value in the Distinct Context of Mobile Technology and Social Media,” *Journal of Service Management*, 24(3), 268–293.

Lau, R.Y.K., Li, C. and Liao, S.S.Y. (2014), “Social Analytics: Learning Fuzzy Product Ontologies for Aspect-Oriented Sentiment Analysis,” *Decision Support Systems*, 65, 80–94.

Lemke, F., Clark, M. and Wilson, H. (2010), “Customer Experience Quality: An Exploration in Business and Consumer Contexts Using Repertory Grid Technique,” *Journal of the Academy of Marketing Science*, 39(6), 846–869.

Lemon, K.N. and Verhoef, P.C. (2016), “Understanding Customer Experience and the Customer Journey,” *Journal of Marketing*, 80(6), 1–62.

Mattila, A.S. and Wirtz, J. (2004), “Consumer Complaining to Firms: The Determinants of Channel Choice,” *Journal of Services Marketing*, 18(2), 147–155.

Maxham, J.G. (2001), “Service Recovery’s Influence on Consumer Satisfaction, Positive Word-Of-Mouth, and Purchase Intentions,” *Journal of Business Research*, 54(1), 11–24.

Meijer, K., Frasincar, F. and Hogenboom, F. (2014), “A Semantic Approach for Extracting Domain Taxonomies from Text,” *Decision Support Systems*, 62, 78–93.

Michel, S., Bowen, D. and Johnston, R. (2009), “Why Service Recovery Fails: Tensions Among Customer, Employee, and Process Perspectives,” *Journal of Service Management*, 20(3), 253 – 273.

van Noort, G. and Willemsen, L.M. (2012), “Online Damage Control: The Effects of Proactive Versus Reactive Webcare Interventions in Consumer-generated and Brand-generated Platforms,” *Journal of Interactive Marketing*, 26(3), 131–140.

Ordenes, F.V., Theodoulidis, B., Burton, J., Gruber, T. and Zaki, M. (2014), “Analyzing Customer Experience Feedback Using Text Mining: A Linguistics-Based Approach,” *Journal of Service Research*, 17(3), 278–295.

Ordenes, F.V., Ludwig, S., de Ruyter, K., Grewal, D. and Wetzels, M. (2017), “Unveiling What Is Written in the Stars: Analyzing Explicit, Implicit and Discourse Patterns of Sentiment in Social Media,” *Journal of Consumer Research*, 43(6), 875-894.

Payne, A. and Frow, P. (2005), “A Strategic Framework for Customer Relationship Management,” *Journal of Marketing*, 69(4), 167–176.

Payne, A.F., Storbacka, K. and Frow, P. (2008), "Managing the Co-creation of Value," *Journal of the Academy of Marketing Science*, 36(1), 83–96.

Prahalad, C.K. and Ramaswamy, V. (2004), "Co-creation Experiences: The Next Practice in Value Creation," *Journal of Interactive Marketing*, 18(3), 5–14.

Rafaeli, A., Ziklik, L. and Doucet, L. (2008), "The Impact of Call Center Employees' Customer Orientation Behaviors on Service Quality," *Journal of Service Research*, 10(3), 239–255.

Ratliff, J.M. and Kunz, M.B. (2014), "A Profile of Social Media Use By Fortune 500 and Inc. 500 Corporations," *Academy of Strategic Management Journal*, 13(2), 61–73.

Rovani, M., Maggi, F.M., de Leoni, M. and van der Aalst, W.M.P. (2015), "Declarative Process Mining in Healthcare," *Expert Systems with Applications*, 42(23), 9236–9251.

Schaefers, T. and Schamari, J. (2016), "Service Recovery via Social Media: The Social Influence Effects of Virtual Presence," *Journal of Service Research*, 19(2), 192–208.

Schoor, C. and Bannert, M. (2012), "Exploring Regulatory Processes During A Computer-Supported Collaborative Learning Task Using Process Mining," *Computers in Human Behavior*, 28(4),1321–1331.

Schweidel, D.A. and Moe, W.W. (2014), "Listening In on Social Media: A Joint Model of Sentiment and Venue Format Choice," *Journal of Marketing Research*, 51(4), 387–402.

Sedrakyan, G., de Weerdt, J. and Snoeck, M. (2016), "Process-Mining Enabled Feedback: "Tell Me What I Did Wrong" v.s. "Tell Me How To Do It Right"," *Computers in Human Behavior*, 57, 352–376.

Simons, J.V. and Kraus, M.E. (2005), "An Analytical Approach for Allocating Service Recovery Efforts to Reduce Internal Failures," *Journal of Service Research*, 7(3), 277–289.

Singh, J.P., Irani, S., Rana, N.P., Dwivedi, Y.K., Saumya, S. and Roy, P.K. (2016), "Predicting the "Helpfulness" of Online Consumer Reviews," *Journal of Business Research*, 70, 346–355

Smith, A. (2013), "The Value Co-destruction Process: A Customer Resource Perspective," *European Journal of Marketing*, 47(11/12),1889–1909.

Smith, A.K., Bolton, R.N. and Wagner, J. (1999), "A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery," *Journal of Marketing Research*, XXXVI, 356–372.

Song, M. and van der Aalst, W.M.P. (2008), "Towards Comprehensive Support for Organizational Mining," *Decision Support Systems*, 46(1), 300–317.

Tirunillai, S. and Tellis, G.J. (2014), "Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation," *Journal of Marketing Research*, 51(4), 463–479.

Ur-Rahman, N. and Harding, J.A. (2012), "Textual Data Mining for Industrial Knowledge Management and Text Classification: A Business Oriented Approach," *Expert Systems with Applications*, 39(5), 4729–4739.

van Vaerenbergh, Y. and Orsingher, C. (2016), "Service Recovery: An Integrative Framework and Research Agenda," *Academy of Management Perspectives*, 30(3), 328–346.

Vázquez-Casielles, R., Álvarez, L.S. and Martín, A.M.D. (2010), "Perceived Justice of Service Recovery Strategies: Impact on Customer Satisfaction and Quality Relationship," *Psychology and Marketing*, 27(5), 487–509.

Verhoef, P.C., Lemon, K.N., Parasuraman, A., Roggeveen, A., Tsiros, M. and Schlesinger, L.A. (2009), “Customer Experience Creation: Determinants, Dynamics and Management Strategies,” *Journal of Retailing*, 85(1), 31–41.

Xu, K., Liao, S.S., Li, J. and Song, Y. (2011), “Mining Comparative Opinions from Customer Reviews for Competitive Intelligence,” *Decision Support Systems*, 50(4), 743–754.

Zeng, D., Chen, H., Lusch, R. and Li, S.-H. (2010), “Social Media Analytics and Intelligence,” *IEEE Intelligent Systems*, 25(6), 13–16.

Chapter 5 A Data-Driven Approach for Supporting Corporate Social Innovation

ABSTRACT

Big data is viewed as “the next big thing in innovation”. Yet, applying a (big) data-driven approach to support social innovation is far less explored compared to such approaches in the areas of technology, information systems, and business. This paper aims to enhance organisations’ performance in corporate social responsibility (CSR) and drive corporate social innovation (CSI) using a data-driven approach. A strategic framework drawing on a value co-creation perspective is proposed and tested by the use of data. The framework clarifies the mechanism that CSI can be promoted through reducing the cognitive distance between a company’s value propositions of CSR and the real stakeholder value. A systematic literature survey was conducted to examine relevant prior research using the framework as a classification scheme. Then, a case study based on a UK retailer was employed to test how CSI based on the data-driven approach could be implemented following the framework. Text-/data-mining techniques were adopted to analyse the retailer’s CSR dialogue data collected from Twitter. The findings from the case study were applied to construct five propositions which deliver testable knowledge. Importantly, the framework and propositions are of practical value as they can serve as operational guidance for practitioners who endeavour to implement data-driven CSI.

Keywords: Corporate Social Responsibility, Social Innovation, Big Data, Text Mining, Social Media, Value Co-creation

5.1 Introduction

Companies increasingly embrace big data and analytics as a new solution to drive innovation (Kiron et al. 2012). The term "big data" describes data with the features of high volume, high variety, and high velocity, known as the "3 Vs" (e.g. McAfee and Brynjolfsson 2012), and two later features – veracity and value – added to form a definition comprising "5 Vs" (e.g. Kunz et al. 2017). The existing innovation models supported by (big) data-driven approaches can be witnessed in using data to improve products/services, digitising physical assets, trading data, combining data within and across industries (e.g. products developed based on the internet of things) and codifying a distinctive service capability (e.g. offering customised pricing) (Parmar et al. 2014). Though data-driven approaches have been widely used in technical and business endeavours, limited empirical research has explored the use of data to address social problems and promote social innovation (Desouza and Smith 2014).

Social problems are often present in the form of wicked problems which are dynamic and complicated as they involve multiple stakeholders with conflicting perspectives (Desouza and Smith 2014). Data can offer easy access to understanding stakeholder expectations towards better CSR practices and emerging market trends. Some applications of using data-driven approaches in addressing social problems have been found in business organisations, such as BT's (British Telecommunications) Better Future Programme that introduces smart meters to collect and analyse user data to help customers reduce their environmental impact. We conceptualise the process of using stakeholder-related data to address social problems and drive corporate social innovation (CSI) as the value co-creation between a company and its stakeholders. Through the co-creation process, companies can facilitate stakeholder knowledge to develop or adjust CSR practices and, thus, enhance social value.

To understand CSI supported by a data-driven approach (defined as data-driven CSI), this research examines social media data. The *Guide to Social Innovation* published by the European Commission (2013, p. 5) identified the importance of social media in transforming the ways citizens and groups relate to public issues and in the speeding up of social innovation. Social media have been viewed as a “public arena of citizenship” (Whelan et al. 2013). On social media, customers and other stakeholders increasingly take on the role of CSR activists and publicly punish companies’ unethical behaviour by spreading negative word of mouth (Korschun and Du 2013). At the same time, companies frequently initiate social chats with stakeholders to spark debates over social issues and broadcast their CSR efforts (Eberle et al. 2013). The diverse CSR activities on social media create a vast volume of dialogue data. The dialogue data is representative of the value co-creation between companies and their stakeholders (Prahalad and Ramaswamy 2004). This is because, through this data, companies can acquire information comprising a mixture of world views of various stakeholders and can thus gain an understanding of the external environment (Lusch and Nambisan 2015).

Analysing social media CSR dialogues to drive CSI still remains unexplored in published research. In particular, social media data is often unstructured data consisting of the majority of big data (e.g. text, image, audio, and video). Desouza and Smith (2014) stressed that data related to social issues is highly grounded in the unstructured format, and few initiatives have approached the use of data to combat existing problems. Compared to structured data (e.g. numerical data), unstructured data often remains unused (Spiess et al. 2014). Today, being data-savvy and employing data analytics to improve innovation have been viewed as a critical approach for companies to obtain competitive advantages (Kiron et al. 2012). However, except for a few large corporations, most of the mid-market firms are struggling to reap data-driven benefits due to their limited understanding and capabilities (Goes 2014).

This research aims to provide a better understanding of the issues of data-driven CSI and offer an operational toolkit as a solution to implementing CSI by the use of data. Several contributions are made by this research. First, this paper offers a framework which adds to the poorly-justified but frequently-used concept of "co-creation" in prior CSI research. The proposed framework contributes to CSI literature by clarifying the mechanism of value co-creation of CSI and specifying the associations between company CSR practices and stakeholder value.

In addition, the framework demonstrates the feasibility of promoting data-driven CSI by highlighting how the cognitive distance can be uncovered and reduced by comparing the differences between a company and its stakeholders. Five propositions are provided as operational guidance that allows practitioners to consider data sources relevant to stakeholder value and understand the cognitive distance disclosed in the data. Finally, this paper demonstrates a data analytical process by facilitating text/data mining to analyse the unstructured social media data. This data analytical approach offers a novel research method and contributes to future studies centred on improving the understanding of data-driven innovation in different research contexts.

5.2 Corporate Social Innovation

The concept of corporate social responsibility (CSR) is an ongoing debate, with over 30 definitions in prior studies (Dahlsrud 2008). More than a management concept, CSR represents an umbrella term comprising several focuses – e.g. four categories of social responsibilities: economic, legal, ethical, and philanthropic, in Carroll's (1979) early corporate social performance model – and importantly, it evolves over time as values change (Carroll 1999). Generally, CSR shows a company's commitment to improving societal well-being by arranging business resources in a way that satisfies the expectations of society (Kotler and Lee 2005, p. 3). Prior research viewed CSR as a strategic tool, allowing companies to obtain broader stakeholder support (Sen et al. 2006), secure customer satisfaction (Luo and Bhattacharya 2006) and improve brand equity and performance (Lai et al. 2010). Also, CSR plays a crucial role in organisational learning and innovation, since CSR programmes allow companies to develop broader and closer relationships with stakeholders and, meanwhile, stimulate knowledge exchange amongst various parties (Luo and Du 2015).

The concept of social innovation has been increasingly adopted to understand the practice of companies creating business value through addressing social problems (Saul 2011). Differing from CSR centred on producing goodwill and strengthening corporate reputation, corporate social innovation attempts to engage a full range of company assets to meet social challenges and create new sources of revenue via a socially relevant innovation system (Mirvis et al. 2016). Social innovation goes beyond CSR and other philanthropic efforts, providing companies with a tangible, direct and near-term approach to creating business impact (Saul 2011). Importantly, social innovation can help businesses overcome the societal barriers hindering their economic growth, such as the aged population or climate change (Mulgan et al. 2007). However, the concept of social innovation is still rather vague, and knowledge of the field remains fragmented (Martinez et al. 2017; van der Have and Rubalcaba 2016).

The European Commission's *Guide to Social Innovation* (2013, p.6) provided a broad definition of social innovation: "the development and implementation of new ideas (products, services, and models) to meet social needs and create new social relationships or collaborations". More specifically, social innovation refers to the design, implementation, and dissemination of new types of social practices in order to promote change to remedy societal problems with respect to the environment, education, employment, culture, health and economic development (Viñals 2013). Although the literature on social innovation has drawn heavily on social entrepreneurship, social enterprises, public sectors and non-profit organisations (e.g. Currie and Seddon 2014; Ionescu and Marga 2015; Viñals 2013), social innovation has long existed in mainstream business corporations (Carberry et al. 2017). Herrera (2015, p. 1469) used the term corporate social innovation (CSI), clarifying social innovation as "a measureable, replicable initiative that uses a new concept or a new application of an existing concept to create shareholder and social value". It is important to note that social innovation and corporate social innovation are grounded in both sociological and economic dimensions, allowing organisations to create social impact and simultaneously obtain economic benefits (Mulgan et al. 2007; Mirvis et al. 2016).

This paper adopts Herrera's (2015) definition of CSI to understand the emerging social value (or social practices) promoted through the collaboration amongst organisations and their stakeholders. CSI is, in its nature, stakeholder-centric, and it can feature as “open innovation”, which highlights that the locus of innovation is shifting away from learning from the past within a single company towards learning in a networking environment where individuals, organisations, and societies exist (Holmes and Moir 2007; Holmes and Smart 2009). Engaging stakeholder knowledge and skills allows companies to search for weak signals of changes and identify new sources of innovation (Holmes and Smart 2009; Risso 2012). Hence, how to manage different knowledge sources and maximise the effect of new knowledge becomes the focal issue of CSI.

This research suggests that a data-driven approach can aid efficient access to and management of diverse knowledge sources within stakeholder-generated data and thereby drive CSI. Desouza and Smith (2014) pointed out that big data is the catalyst for social innovation. They offered several suggestions to discover the insights from data in informing decision-making and solving the world’s toughest social problems, such as building data banks on critical issues, engaging stakeholders (e.g. citizens) in idea-creation activities, building a cadre of data curators and analysts, and providing virtual experimentation platforms (Desouza and Smith 2014). Our work adds to data-driven social innovation by exploring the potential of data for managing stakeholder value with respect to the acquisition of representative stakeholder data, knowledge discovery using data analytics, and identification of critical insights for decision-making.

5.3 Research Methodology

A two-phase research methodology was adopted in this paper. In the first phase, a systematic literature survey of journal articles dealing with “corporate social innovation” (CSI) was conducted. According to Altuna et al. (2015), prior CSR research was limited in the extent to which it examined how for-profit organisations can manage social innovation. The findings from the first phase present an extensive review of relevant studies on the defined CSI classification scheme and indicate the gaps in prior research. In the second phase of research, a case study based on a UK grocery retailer was conducted to test the framework in the real-world setting. This case study aimed to investigate data-driven CSI by analysing the company's published CSR data and the dialogue-based data collected from the company's Twitter platform. The case study provides a deductive approach, producing testable propositions to improve the understanding of data-driven CSI.

5.4 Phase I: Systematic Literature Survey of Journal Articles

A comprehensive literature survey was conducted on a similar approach suggested by Fosso Wamba et al. (2015) and Ngai et al. (2009). This approach consists of three activities: (i) developing a classification framework, (ii) conducting the literature review, and (iii) realising the classification of relevant journal articles (Fosso Wamba et al. 2015). The first activity is discussed in **Section 5.4.1**, and a framework is codified for the classification of CSI literature. In **Section 5.4.2**, the second and third activities are discussed.

5.4.1 The Classification Framework

In developing a classification framework, we consulted the works of Herrera (2015) and Mirvis et al. (2016) to identify the mechanism and the relevant elements of CSI in prior research. A visualisation of the classification framework is provided in Figure 5.1.

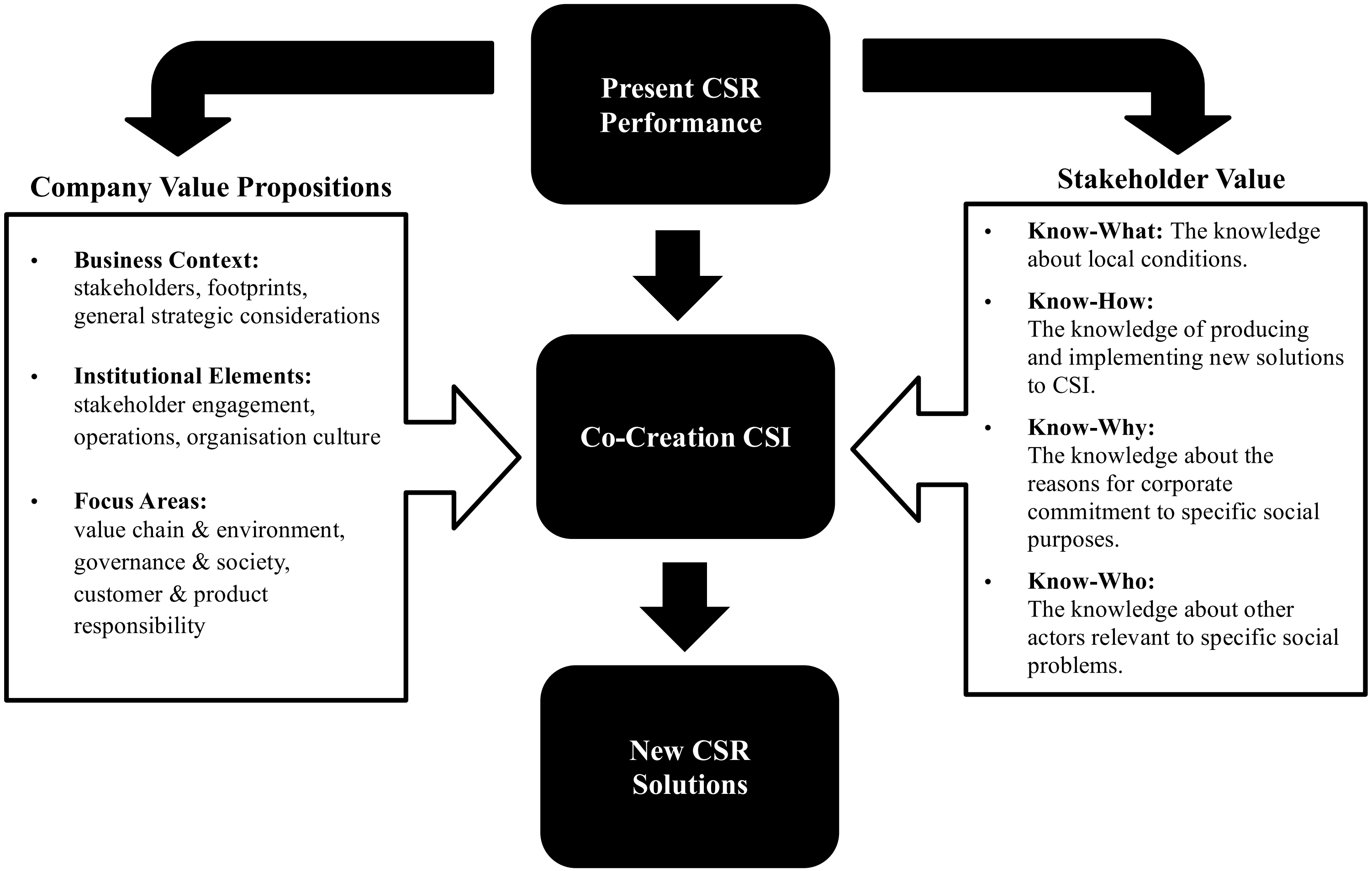

Figure 5.1 The Classification Framework

Mirvis et al. (2016) explained that the mechanism of CSI works on the deeper cooperation through which a company and its stakeholders are better able to co-create a new sustainable solution to specific social ills and achieve shared value. According to Prahalad and Ramaswamy (2004), co-creation is the basis for value creation and it is the key to uncover new sources of competitive advantages. Co-creation CSI can be interpreted as a co-creation process of enhancing social value and it serves as a transformation from existing CSR solutions to new CSR solutions . A company's CSR activities can only be seen as value propositions that connect different social systems such as consumers, suppliers, and other stakeholders and potentially improve the well-being of the systems (Vargo et al. 2008). The real value is realised when stakeholders perceive that social value has improved through better CSR value propositions.

Often, CSR activities are viewed differently by various stakeholders at different times (Campbell 2007). This is because "people will perceive, interpret, understand and evaluate the world differently to the extent that they have constructed their cognition along different, weakly connected life paths" (Nooteboom 2009, pp. 66–67). Cognitive distance describes the difference in knowledge, skills or cognitive frames amongst actors (Hendriks-Jansen 1996). Le Ber and Branzei (2010b) specified the process whereby the cognitive distance between a company and its stakeholders is reduced as value frame fusion. Lusch and Nambisan (2015) stressed that identifying and reducing cognitive distance is key to co-created innovation. In this framework, we highlight that reducing cognitive distance and obtaining shared value allow for the development of new services and products (value propositions) that are more compelling for stakeholders' expectations (Herrera 2015; Mirvis et al. 2016).

Co-creation CSI involves two main parties – the company and its stakeholders – consisting of many sub-groups, such as consumers, communities, suppliers, and competitors. Both parties enter into the co-creation CSI after recognising the necessity of improving the present CSR performance of the focal company. When the CSR performance is interpreted differently in the company's value propositions and stakeholders' value, the cognitive distance occurs. The key elements of the company's value proposition of CSR are evaluated using Herrera's (2015) framework, including business context, institutional elements, and focus areas.
framework, including business context, institutional elements, and focus areas.

(i) *Business Context* focuses on how enterprises can integrate market/non-market conditions into strategic processes and includes the components such as stakeholders, corporate footprints (e.g. value chain of the company that creates economic, environmental, and social impact), and general strategic considerations (e.g. the company's core value, philosophies, resources, and competences).

(ii) *Institutional elements* consist of stakeholder engagement, operations policies, structures and processes, and organisational culture. Notably, the institutional elements are used to drive, enable and embed the company's value and processes into specific CSR focus areas.

(iii) *Focus areas* indicate the specific CSR dimensions or social goals that companies can enhance through better management of company resources. The CSR focus areas are broken down into three aspects: governance and society, customer and product responsibility, and value chain and environment.

In contrast, Mirvis et al.'s (2016) knowledge exchange model was used to interpret stakeholder value, which is explained by four types of stakeholder knowledge: know-what, know-how, know-who, and know-why.

(i) *Know-what* is the knowledge about "facts", which is close to the meaning of information (Lundvall and Johnson 2016, p.112). Know-what knowledge has been discussed in prior organisation learning research, referring to how an organisation best arranges itself regarding the facts of business, market, distribution or sales (Mcelroy 2003, p. 46). In the CSI context, Mirvis et al. (2016) conceptualised stakeholder know-what as the knowledge of local conditions where the supply chain or target market exists.

(ii) *Know-how* is a learning-by-doing process and it accumulates with experience over time (Garud 1997). Know-how knowledge refers to the crucial skills required to trigger economic growth through knowing how to form new methods and new products/services and how to develop new technologies (Lundvall and Johnson 2016, p. 113). Thus, the stakeholder know-how in the CSI context is related to the knowledge of producing and implementing social innovation in the high-uncertainty environment (Mirvis et al. 2016).

(iii) *Know-why* refers to the knowledge of principles and laws that contain a deep understanding of causal relationships and ambiguous associations with observed things (Lundvall and Johnson 2016, p. 112). Know-why knowledge is established on know-what and know-how, revealing the contextual factors associated with the former two. In the CSI context, stakeholders' know-why is related to their understanding of corporate commitment to specific social purposes, such as environmental issues (Mirvis et al. 2016).

(iv) *Know-who* is people-related knowledge that reveals information about who knows what and who knows what to do in or outside the company or the company's social networks (Johnson et al. 2002). Mirvis et al. (2016) stated that stakeholder know-who identifies important social ties in the knowledge acquisition and transfer network based on a shared social goal.

5.4.2 Literature Review Search Strategies

On the basis of the classification framework, the literature review was conducted, and relevant articles were classified accordingly. A search within the timeframe of 2003 to 2016 was considered to be representative of the take-off phase of social innovation research (van der Have and Rubalcaba 2016). Notably, most of the research on social innovation tended to focus on the non-profits, resulting in a limited understanding of CSI in business organisations (Altuna et al. 2015). To advance the area of concern, the literature survey was implemented to explore CSI in business organisations rather than the public sector, non-profits, and non-governmental organisations. Also, we excluded the articles centred on social entrepreneurs and social enterprises, considering that their business philosophies, purposes, and priorities are different from general for-profit organisations.

The literature searches with the keywords "corporate social innovation", "social innovation" AND "corporate social responsibility", and "social innovation" AND "business(es)" were conducted within two scholarly research databases in the business and economic domain: ABI/Inform Global and Business Source Premier. The initial searches resulted in 773 articles. We screened the titles, abstracts, keywords, and references of all articles, excluding the articles written in non-English, book reviews, editorial articles, and commentaries. In the next step, we merged the resulting datasets from ABI/Inform Global and Business Source Premier, removing the duplicated articles. The final dataset for further analysis included 32 articles. After reviewing the articles, we categorised these articles following the classification framework.

5.5 Phase II: Case Study of a UK Grocery Retailer

In the second phase, a case study based on a UK grocery retailer was conducted to explore how a data-driven approach can help the implementation of co-creation CSI. When it comes to theory-building, the case study approach is suitable for understanding the chaotic phenomenon by conceptualising the emerging body of knowledge (Yin 2003). The company embraces a wide range of CSR activities, including the community as its priority, reducing environmental impacts, and promoting ethical trade. More than 50 key performance indicators (KPIs) are set by the company to monitor their CSR activities. Hence, the company is considered as a suitable research subject.

The proposed framework (see Figure 5.1) was applied to the analysis of the company's CSR with respect to business context, institutional elements, and focus areas, and the stakeholder value regarding know-what, know-how, know-who, and know-why. As stressed earlier, CSI is driven by a co-creation process through which the cognitive distance between a company and its stakeholders can be identified. In the case study, co-creation CSI was examined by comparing the consistency between the company's CSR-related data and the stakeholder-generated data.

5.5.1 Data Collection Methods

Multiple sources of data were used to investigate how a data-driven approach could be developed and applied to support the co-creation CSI of the company. The considered data included data regarding the company's CSR practices and data containing the stakeholder value. The first type of data involved the company's published CSR reports, annual reports, and newsletters. The second type of data could be collected by traditional research methods such as surveys or interviews with stakeholders or by examining new sources of big data. Korschun and Du (2013) stated that more and more companies initiate "virtual CSR dialogues" via social media to promote co-creation CSR. Within these virtual dialogues, both the company's CSR activities and the stakeholder value are embedded. We observed the virtual dialogues taking place on several social media platforms of the company and chose CSR dialogues on Twitter for further analysis.

As this research specifically focuses on the (big) data-driven approach in supporting co-creation CSI, the data analysis of the case study was mainly centred on Twitter dialogue data (known as "tweets") to identify the cognitive distance and find the opportunities for CSI.

5.5.2 Research Setting: Analysing Co-Creation CSI within Social Media Data

According to Roberts and Piller (2016), the mainstream of using social media to drive innovation can be classified into three activities: exploring, co-creating and communicating. Exploring activities are conducted to identify new market trends and reveal critical insights embedded in the user-generated content, such as Twitter feeds or Facebook postings. Co-creating activities are initiated by companies to collaborate with critical actors in the innovation processes. Finally, communicating activities aim to create dialogues with audiences and stimulate interests via social media.

We observed the Twitter dialogues of the company, analysing how the three innovation activities on social media were implemented in the context of CSI. The co-creating dialogues were initiated by the company to engage Twitter actors in the design of CSR programmes such as raising CSR issues, suggesting the allocation of funds, and offering specific solutions to CSR practices. The communicating dialogues could be initiated by either the company or other Twitter users. The main purpose was to turn Twitter users into brand ambassadors who spread positive views of the company's CSR efforts via word of mouth. Through communicating dialogues, the company could understand which CSR activities are more impactful for triggering stakeholders' positive feelings. Finally, the exploring activity on dialogues initiated by Twitter users to attract the company's attention on special social problems could unveil the emerging CSR trends, competitive intelligence (e.g. competitors' CSR strengths) and dissatisfactory CSR practices of the company.

To investigate CSI dialogues, we developed an analytical process using text mining and data mining. Text mining was employed to enable speedy and automatic information extraction, which is a process that used to be conducted manually (Ordenes et al. 2014). Text mining aims to uncover hidden trends, patterns or rules from the sizable text-based data, transforming the unstructured data into a structured format (He et al. 2013). Three types of text mining were often used in prior research: linguistic, statistical, and a hybrid approach (Meijer et al. 2014). Linguistic methods rely on natural language processing (NLP), such as tokenization, stemming, part-of-speech (POS) tagging and sentiment tagging, for term extraction (Xu et al. 2011). Linguistic methods require training instances to define text extraction rules. In contrast, statistical approaches extract information using statistics measures. The widely used statistical methods include term frequency (TF) and topic modelling (Aggarwal and Zhai 2013, p. 5). Hybrid methods combine the former approaches by calculating scores of the extracted terms after linguistic processing (Meijer et al. 2014). In the case study, both linguistic and statistical methods were employed to extract hidden insights in Twitter dialogues.

Figure 5.2 shows the data analytical process of the case study. Firstly, the company's CSR-related documents (e.g. CSR reports, annual reports, and newsletter) were manually analysed to define the boundary of the company's CSR practices. Then, the linguistic-based text mining was applied to identify and classify Twitter dialogues related to CSR, tweet sentiment tagging and named entity tagging. The linguistic-based extraction required prior knowledge, which was acquired by examining the company's CSR documents to construct a dictionary and using external lexical resources. Subsequently, the statistical methods were used to identify the unknown information mentioned by stakeholders in the Twitter dialogues.

Though text mining places emphasis on automatic information extraction, high levels of human expert involvement are often required for text rule indexing and data labelling. In order to identify the cognitive distance between the company and its stakeholders, expert workshops were organised to analyse the stakeholder tweets and label the tweets into new CSI ideas, improvement CSI ideas, and non-ideas. Finally, based on the results of text mining and the idea labels, we performed experiments using the decision tree technique to unveil critical rules of CSI ideas. More details of the data analysis process are explained in the following sections.

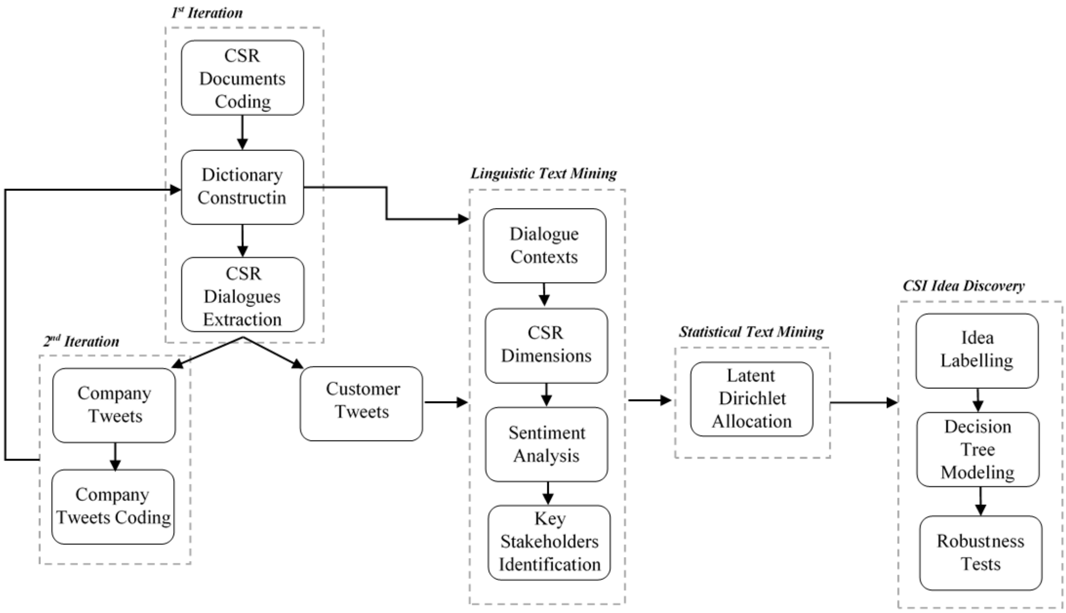

Figure 5.2 Data Analysis Process of the Case Study

5.5.3 Thematic Coding of the Company's CSR Documents

To understand the company's CSR activities in practice, two domain experts analysed the company's CSR reports, annual reports and newsletters published in the past five years (2011–2015). We adopted template analysis to code relevant topics following ISO 26000, which is an international standard of CSR that has been widely adopted in prior CSR studies (e.g. Castka and Balzarova 2008; Helms et al. 2012). ISO 26000 covers six main CSR dimensions (*Human Rights, Labour Practices, The Environment, Fair Operating Practices, Consumer Issues, and Community Involvement*) and 36 sub-dimensions (ISO 26000 2010). The manually extracted terms were coded and classified into the six main CSR dimensions.

5.5.4 Text Mining of Twitter Dialogue Data

The Twitter dataset includes 70,800 Twitter posts of the company and its stakeholders over a six-month period (7,201 dialogues). Text-mining approach was applied to examine the company's CSR practices in company tweets and stakeholder knowledge regarding know-what, know-who, and know-how within stakeholder tweets. In terms of stakeholder know-why, this knowledge holds stakeholders' reasoning about the corporate commitment to social purposes (Mirvis et al. 2016). This information could not be obtained by purely mining stakeholder tweets in the dialogues but rather by the further investigation of multiple observed phenomena (factors) to reveal their hidden associations or causal relationships. We will discuss the stakeholder know-why extraction in **Section 5.5.5**.

Identification of CSR Dialogues and CSI Contexts

Since the Twitter dataset contained both CSR and non-CSR dialogues, the first task was to identify the CSR dialogues. A dictionary was constructed for CSR dialogue extraction. This approach requires a manually tagged training sample as seed words for information extraction (AlSumait et al. 2010). We applied the results of thematic coding from the company's CSR documents to build the primary dictionary (first iteration) for automated information extraction from the Twitter dataset.

Notably, the initial dictionary was built on the company's self-disclosed reports and newsletter, and thus it might have limitations in accurately identifying the CSR tweets due to the platform's unique communication culture in using hashtags and abbreviations. To address this, the second iteration was conducted using company tweets that had been recognised as CSR dialogues. In addition to coding the new terms in the company tweets, we further coded terms demonstrating different CSI contexts: exploring, co-creating, and communicating activities (Roberts and Piller 2016). Table 5.1 provides examples of Tweet dialogue coding. The fine-tuned dictionary was applied to conduct the second extraction of CSR dialogue and to identify the CSI contexts.

Table 5.1 Examples of Tweet CSR Coding

Tweet Content	CSR Dimensions	Dialogue Context
Who should we partner with to champion change? We're looking for charity partnership so vote now (URL)	charity partnership *(Community Involvement)*	vote now *(Co-creating Dialogue)*
Juan is a #Fairtrade wine producer in Argentina. Visit the # (CSR campaign) website to speak to him directly. (URL)	#Fairtrade *(Fair Operating Practices)*	# (CSR campaign) website *(Communicating Dialogue)*
@User ID You can find out more about our commitment to animal welfare here: (URL)	animal welfare *(The Environment)*	@User ID *(Exploring Dialogue)*

Know-What Extraction

Stakeholder know-what relates to the knowledge of local conditions crucial to the company's business operations. To extract stakeholder know-what, we separated the stakeholder tweets from the identified CSR dialogues and obtained 3,468 tweets. Again, the dictionary was used to examine stakeholder tweets and extract the six CSR dimensions mentioned by stakeholders.

Know-Who Extraction

Stakeholder know-who refers to the knowledge of important social ties and stakeholders of the company. We screened business partners (e.g. charities, non-profits) and competitors of the company and built up a list for named entity tagging of stakeholder tweets. Moreover, the user names mentioned in tweets with a high frequency were examined and classified into the named entity list. As such, we could extract the critical stakeholders mentioned in tweets.

Know-How Extraction

Know-how knowledge is a learning-by-doing process and it accumulates with experience over time (Garud 1997). Stakeholders' know-how indicates their learning and experiences about the ideal CSR practices that should be implemented by the company. To understand stakeholder know-how, we first examined how they felt about the company's CSR practice and then examined their specific comments on that practice. In addition to the company's CSR practice, we also assessed the CSR practice not considered by the company but mentioned in stakeholders' tweets.

Sentiment analysis is often used by practitioners to examine the polarity of a document, revealing how people feel about a topic, such as customer opinions toward a product/brand (Xu et al. 2011). Numerous commonly shared sentiment lexicons have been used in Twitter opinion-mining research, such as SentiWordNet 3.0 (available from http://sentiwordnet.isti.cnr.it). In this paper, we analysed tweet sentiment using SentiWordNet 3.0, SentiStrength and the method proposed by Kolchyna et al. (2016) that added emoticons into the sentiment lexicon. Based on the sentiment lexicons, each tweet was categorised into positive, negative or neutral. We decided upon the final tweet sentiment by assessing the agreement between the three sentiment lexicons. For a conflicting tweet sentiment, human experts examined the tweet and assigned it with a suitable sentiment label.

After identifying sentiment within each of the stakeholder tweets, we applied a topic-mining technique to recognise and classify the unknown knowledge within the tweets, such as specific comments about the company's CSR practices or suggestions on the unimplemented CSR practice. For example, a tweet referring to "please support local framers in Yorkshire" may be assigned to the topic "sourcing location". Unlike the linguistic-based methods enabling extraction of well-defined knowledge (prior knowledge), topic modelling facilitates statistical methods that examine the undefined concepts in tweets.

Latent Dirichlet allocation (LDA), one of the best topic-modelling techniques, was used to discover unknown information in the stakeholder tweets. LDA uses the Bayesian learning algorithm to extract context-specific features (Tirunillai and Tellis 2014). In LDA processing, a document is transformed into a "bag-of-words", with the words associated with particular topics (Wei and Croft 2006). Each document is then presented by a mixture of latent topics with different probabilities, and the topic with the highest probability will be assigned to the document. Before conducting LDA on the tweets, we removed stop words (e.g. "the", "and") that are used for connection and grammar yet have trivial meanings (Tirunillai and Tellis 2014) and words that had been labelled in the CSR dictionary. In this way, the known information can be excluded to allow unknown information to be unearthed.

The LDA model requires researchers to specify the number of topics and words in a topic being generated. The volume of stakeholder tweets for LDA modelling was only 3,468 tweets, and thus a smaller number of topics was used: the topic number was set from 10 to 50 topics with an interval of 10 (Lo et al. 2015). We selected the number of words in a topic as four words, considering that the average word count of each tweet was 4.18 after removing stop words and labelled words. Finally, we tagged each topic based on the extracted terms to summarise the represented meanings.

5.5.5 CSI Idea Discovery

5.5.5.1 Stakeholder Tweets Labelling

In **Section 5.5.4**, we extracted the information regarding the company's CSR practices and stakeholder knowledge. This helped us to understand the cognitive distance between the company and its stakeholders and to learn if a stakeholder tweet contained a CSI idea enabling the company to improve the CSR performance. We organised expert workshops to collectively label 3,468 stakeholder tweets into non-ideas, improvement CSI ideas and new CSI ideas. Six coders participated in the four-hour workshops, receiving a fixed payment.

According to Thorleuchter et al. (2010), an idea should consist of a problem and solutions. Based on this principle, we defined a non-idea as a tweet containing either no problem/solution or a known problem already tackled by the company. An improvement CSI idea referred to a known problem/solution for which the company had poor performance. Finally, a new idea represented an unknown problem or a new solution for a known problem. Table 5.2 provides examples of the idea labelling.

Table 5.2 Examples of CSI Idea Labelling

Stakeholder Tweets	The Company's CSR Activities	Emerging CSR Issues	Label
@Study Company do you do charity bucket collects/bag packs in store??	charity bucket	-	Non-Ideas
@Study Company I'm not bothered by wonky veg. I bet millions of others feel the same! help stop food waste	food waste	wonky veg *(Defined issue in CSR reports)*	Improvement Ideas
@Study Company ethical? You're shutting #pubs and tearing communities apart. Don't make us leave!!!	communities	shutting #pubs *(Undefined issue in CSR reports)*	New Ideas

In the workshop, each tweet was labelled three times for comparing the agreement amongst coders. If at least two of the labels reached the agreement, the idea label could be finalised. Otherwise, tweets possessing three different idea labels were classified as suspended. This approach was suggested by Kuehl et al. (2016). As a result, we defined 2,480 tweets as non-ideas, 763 tweets as improvement ideas, 119 tweets as new ideas and 106 tweets as suspended.

5.5.5.2 Decision Tree Modelling (Know-Why Extraction)

We used decision tree modelling to disclose the associations amongst the text-mining resulting variables related to know-what, know-who, and know-how and their contributions to CSI ideas (e.g. Kuehl et al. 2016). As shown in Table 5.3, the predictor variables were the stakeholder knowledge generated in the text-mining stage, and the response variable was the "CSI idea" identified in the expert labelling workshops.

Decision tree modelling describes sequences of interrelated decisions, classifying entities and making predictions (Ngai et al. 2009). The results of a decision tree model are manifested in a hierarchical structure with three types of node: the 'internal' node denotes different levels of independent variable splitting into homogeneous classes, the 'leaf' nodes represent the outcome of splitting, and the 'root' node is a conditional attribute (Chien and Chen 2008). Moreover, the results can be interpreted based on conditional rules, such as "*if* Competitor X is mentioned in tweets (know-who), then the environment issues (know-what) are more likely to be perceived negatively (know-how)".

As a supervised machine learning technique, decision tree modelling requires a training sample to induce rules. The performance of models in the training sample is then evaluated in a test sample. Several decision tree algorithms can be used for tree pruning, including C5.0, classification and regression trees (CART) and chi-squared automatic interaction detection (CHAID). In the case study, the C5.0 algorithm was chosen due to its robustness (Duchessi and Lauría 2013).

It is worth noting that tweets labelled as CSI idea (28%) were outnumbered by the non-idea tweets (72%). Modelling an imbalanced dataset can be a challenge, as algorithms are often biased towards the majority class, which hampers the predictability of rare events. However, in our case, despite the presence of a data imbalance issue, we found that decision tree modelling performed well. Hence, the dataset remained imbalanced without re-balancing.

Table 5.3 Variables for Decision Tree Modelling

	Variables	Data Type
Predictor Variables		
Stakeholder Know-What	Community Involvement	Dichotomous
	The Environment	Dichotomous
	Employee Practices	Dichotomous
	Consumer Issues	Dichotomous
	Fair Operating Practices	Dichotomous
	Human Rights	Dichotomous
Stakeholder Know-Who	Competitors	Dichotomous
	External Partners	Dichotomous
Stakeholder Know-How	Tweet Sentiment	Nominal
	LDA Topics (10-50 topics)	Numerical
Response Variable	CSI Ideas	Nominal

5.5.6 Robustness Tests

To evaluate the performance of the decision tree models, we applied the widely-used performance measures: precision, recall, and F-measure (Sokolova and Lapalme 2009). We ran validation experiments to alleviate the problem of over-fitting (Abrahams et al. 2013). Two potential over-fitting problems were examined, including the proportion of the training sample and the number of LDA topics used for modelling. To assess the performance, we first varied the proportion of the training sample from 10% to 90%, in 10% increments. Then, we selected the best performance proportion of the training/testing sample and examined the number of topics, from 0 to 50 topics, in 10-topic increments.

5.6 Results

5.6.1 Findings from the Literature Survey

This section presents and discusses the findings from the review of past journal articles dealing with "corporate social innovation". Table 5.4 provides the distribution of the examined 32 articles following the classification framework with three main dimensions: co-creation CSI, company's value propositions of CSR, and stakeholder value. Most of the articles cover more than one dimension and sub-dimension.

First, we can notice that there are only nine articles (28.12%) embracing the concepts of "co-creation" or "value co-creation" to interpret the CSI process promoted by the engagement with stakeholders. Though the collaboration with stakeholders has been highlighted in most of the examined articles, the concept of value co-creation is barely investigated. In the nine articles, five focus on the cross-sector partnerships in which a company collaborates with non-profits to obtain crucial external knowledge and co-create social value.

Regarding the dimension of the company's value propositions, three sub-dimensions are specified: business context, institutional elements, and CSR focus areas. Each sub-dimension also contains three elements. As shown in Table 5.4, within the business context, 25 articles (78.13%) relate to “Stakeholder”, 17 articles (53.12%) relate to “General Strategic Considerations”, and only nine articles (28.12%) examine “Footprint”. In terms of institutional elements, “Stakeholder Engagement” (21 articles, 65.63%) and “Operations” (24 articles, 75.00%) attract more researcher interest than “Organisation Culture” (10 articles, 31.25%). Finally, the CSR focus areas of “Governance & Society” (17 articles, 53.12%) and “Value Chain & Environment” (12 articles, 37.50%) are more popular in prior CSI research. In contrast, the area of “Customer & Product Responsibility” only contains seven articles (21.88%).

It is important to note that the main dimension of stakeholder value proposed in the framework has only been examined to a limited extent in previous CSI works. We found more prior research on “Know-How” (eight articles, 25.00%) and “Know-What” (six articles, 18.75%) than the research on stakeholder “Know-Who” (two articles, 6.25%) and “Know-Why” (one article, 6.25%). In the remainder of articles relevant to stakeholder value, the authors, instead of specifying particular types of stakeholder knowledge, focus on the general knowledge in their research.

The results from the literature survey reveal that co-creation CSI and stakeholder value regarding particular types of knowledge remain far from being thoroughly investigated. To improve the understanding of these research areas, a case study was conducted, and the results are discussed in the following section.

Table 5.4 Distribution of Articles According to the Classification Framework

Dimension	Sub-dimension	References	#	%
Co-creation CSI		Le Ber and Branzei (2010a, 2010b); To (2016); Herrera (2016); Murphy et al. (2012); Pressentin (2017); Mattera and Baena (2015); Schweitzer et al. (2015); Holmes and Smart (2009)	9	28.12%
Company's Value Propositions of CSR	Business Context			
	- Stakeholder	Carberry et al. (2017); Le Ber and Branzei (2010); Linna (2012); To (2016); Muthuri et al. (2012); Kolk and Lenfant (2015); Risso (2012); Herrera (2016); Murphy et al. (2012); Selsky and Parker (2010); Salim Saji and Ellingstad (2016); Warnecke (2017); Battisti (2012); Furmańska-Maruszak and Sudolska (2016); Olejniczuk-Merta (2015); Segarra-Oña et al. (2017); Manning and Roessler (2014); Seitanidi (2008); Hanke and Stark (2009); Le Ber and Branzei (2010b); Altuna et al. (2015); Mattera and Baena (2015); Schweitzer et al. (2015); Holmes and Smart (2009); Holmes and Moir (2007)	25	78.13%
	- Footprints	Linna (2012); Saul (2011); Abaza (2017); Herrera (2016); Selsky and Parker (2010); Furmańska-Maruszak and Sudolska (2016); Olejniczuk-Merta (2015); Harazin and Kósi (2013); Mattera and Baena (2015)	9	28.12%

- General Strategic Considerations	Carberry et al. (2017); Le Ber and Branzei (2010); Witell et al. (2016); Saul (2011); To (2016); Muthuri et al. (2012); Kolk and Lenfant (2015); Maiolini et al. (2016); Risso (2012); Herrera (2016); Pressentin (2017); Salim Saji and Ellingstad (2016); Warnecke (2017); Battisti (2012); Manning and Roessler (2014); Seitanidi (2008); Groot and Dankbaar (2014)	17	53.12%
Institutional Elements			
- Stakeholder Engagement	Le Ber and Branzei (2010a, 2010b); To (2016); Muthuri et al. (2012); Kolk and Lenfant (2015); Risso (2012); Herrera (2016); Murphy et al. (2012); Selsky and Parker (2010); Salim Saji and Ellingstad (2016); Battisti (2012); Olejniczuk-Merta (2015); Segarra-Oña et al. (2017); Manning and Roessler (2014); Seitanidi (2008); Hanke and Stark (2009); Altuna et al. (2015); Mattera and Baena (2015); Schweitzer et al. (2015); Holmes and Smart (2009); Holmes and Moir (2007)	21	65.63%
- Operation	Carberry et al. (2017); Le Ber and Branzei (2010); Witell et al. (2016); Linna (2012); Saul (2011); To (2016); Abaza (2017); Muthuri et al. (2012); Maiolini et al. (2016); Risso (2012); Herrera (2016); Murphy et al. (2012); Selsky and Parker (2010); Salim Saji and Ellingstad (2016); Warnecke (2017); Battisti (2012); Olejniczuk-Merta (2015); Segarra-Oña et al. (2017); Manning and Roessler (2014); Seitanidi (2008); Hanke and Stark (2009); Harazin and Kósi (2013); Groot and Dankbaar (2014); Altuna et al. (2015)	24	75.00%
- Organisation culture	Le Ber and Branzei (2010a, 2010b); Saul (2011); Muthuri et al.	10	31.25%

		(2012); Kolk and Lenfant (2015); Herrera (2016); Pressentin (2017); Warnecke (2017); Furmańska-Maruszak and Sudolska (2016); Groot and Dankbaar (2014)		
	CRS Focus Areas			
	- Value Chain & Environment	Carberry et al. (2017); Witell et al. (2016); Saul (2011); To (2016); Abaza (2017); Muthuri et al. (2012); Kolk and Lenfant (2015); Maiolini et al. (2016); Herrera (2016); Segarra-Oña et al. (2017); Harazin and Kósi (2013); Holmes and Smart (2009)	12	37.50%
	- Governance & Society	Le Ber and Branzei (2010a, 2010b); Witell et al. (2016); Saul (2011); Maiolini et al. (2016); Risso (2012); Herrera (2016); Murphy et al. (2012); Salim Saji and Ellingstad (2016); Warnecke (2017); Furmańska-Maruszak and Sudolska (2016); Segarra-Oña et al. (2017); Manning and Roessler (2014); Seitanidi (2008); Harazin and Kósi (2013); Altuna et al. (2015); Schweitzer et al. (2015)	17	53.12%
	- Customer & Product Responsibility	Witell et al. (2016); Linna (2012); Saul (2011); Maiolini et al. (2016); Battisti (2012); Olejniczuk-Merta (2015); Harazin and Kósi (2013)	7	21.88%
Stakeholder Value	Know-What	Carberry et al. (2017); To (2016); Muthuri et al. (2012); Battisti (2012); Manning and Roessler (2014); Altuna et al. (2015)	6	18.75%
	Know-How	To (2016); Muthuri et al. (2012); Kolk and Lenfant (2015); Risso (2012); Salim Saji and Ellingstad (2016); Mattera and Baena (2015); Manning and Roessler (2014); Battisti (2012)	8	25.00%
	Know-Who	To (2016); Mattera and Baena (2015);	2	6.25%

Know-Why	To (2016)	1	3.13%
General knowledge	Linna (2012); Herrera (2016); Murphy et al. (2012); Warnecke (2017); Segarra-Oña et al. (2017); Le Ber and Branzei (2010b); Schweitzer et al. (2015); Holmes and Smart (2009); Holmes and Moir (2007)	9	28.12%
Total		**32**	**100%**

5.6.2 Findings from the Case Study

5.6.2.1 Text-Mining Results

In the case study, we analysed the company's CSR practices and stakeholder value in terms of stakeholder know-what, know-who and know-how in the three CSI contexts: exploring, co-creating, and communicating. Comparisons were made regarding how the three types of stakeholder knowledge vary in the different CSI contexts. The results are provided in Figures 5.3–5.6.

Exploring Dialogue Context

Exploring context focuses on insight discovery in stakeholder-initiated tweets that serve as open-script dialogues. The exploring context is found to host more diverse information compared to the co-creating and communicating context. Figure 5.3 shows that *Community Involvement* (40%) is the most popular CSR dimension mentioned by stakeholders. It is consistent with the company's CSR strategies – community as the priority highlighted in its CSR documents. The results imply that stakeholder know-what in the exploring context is consistent with the company's main focus of CSR.

As for stakeholder know-who, the exploring context contains more competitor engagement tweets than the other two contexts. This may indicate an emerging market trend to which the company should pay more attention. For example, we found numerous stakeholder tweets mentioning the new living wage guidance published by the UK government. Stakeholders questioned the retailer about its employee payment and particularly related the issue to one of the company's competitors who had agreed to raise the payment. The tweets stated:

Well done @Competitor for deciding to pay all their staff the full living wage. Come on @The Company do the same.

In terms of stakeholder know-how, about 50% of tweets present a positive sentiment. When examining the sentiment in each CSR dimension, we found that *Labour Practices* and *Consumer Issues* contain more negative sentiment tweets (60.76% and 49.17%, respectively) than positive and neutral tweets. This suggests that stakeholders were less satisfied with these two CSR dimensions. We further examined the relevant topics of *Consumer Issues* to understand the associated problems. Table 5.5 details the word list of 10 topics extracted by LDA. These topics were assigned with tags referring to specific products, information, or CSR competition & campaign that were perceived or experienced by stakeholders. One of the relevant topics to *Consumer Issues* is Topic 10, which contains the terms "bacon", "suffolk", "dutch", and "suggest". When further examining the tweets within Topic 10, we found that this topic pointed to an issue about a misleading label:

Another cracker from @The Company. Bit misleading suggesting this is local suffolk bacon...(it's) Dutch.

Co-Creating Dialogue Context

The co-creating context is often initiated by the company to engage stakeholders in the design of CSR programmes. As shown in Figure 5.3, stakeholder know-what in the co-creating context is mainly centred on *Community Involvement*, and 50% of the tweets are associated with charity partners (see Figure 5.4).

Stakeholders indicate their know-who by highlighting the critical local groups or charities needed for the company's support. For example, a consumer group posted tweets seeking financial aid for a local fund competition:

Voting starts 2 days for @The Company local fund. Pls vote for @Charity Partner's lifetime commitment of support to complex disabilities.

Also, the names of local groups and charities participating in the fund competition are uncovered in stakeholder know-how. As can be seen in Figure 5.6, Topic 6 was found to be the most relevant to the co-creating context. In Table 5.5, Topic 6 also discloses the names of active charities and local organisations (e.g. "shakespeare," "ayrshire") as well as their status in the fund competition (e.g. "shortlist").

Communicating Dialogue Context

The communicating context is aimed at raising stakeholder awareness about the company's CSR efforts and triggering public interest in sharing and retweeting information. Therefore, most of the tweets in this context present a positive sentiment (see Figure 5.5). Stakeholder know-who is especially important, with 70% of the stakeholder tweets containing the names of charity partners (see Figure 5.4). In addition to sharing the company's CSR activities, stakeholders also shared CSR campaigns launched by third parties (e.g. charities) who engaged with the company's resources. For example, a partner charity launched a fair trade campaign by using the company's voucher and tweeted:

WIN a £100 @The Company voucher, RT click here: URL; follow the instructions #CampaignName

In this way, it boosted the number of tweets regarding *Fair Operating Practices* due to retweeting and sharing information (stakeholder know-what, see Figure 5.3). This is also revealed in stakeholder know-how. The highest proportion of communicating tweets is found in Topic 9, with related terms such as "win" and "£100" (see Figure 5.6 and Table 5.5).

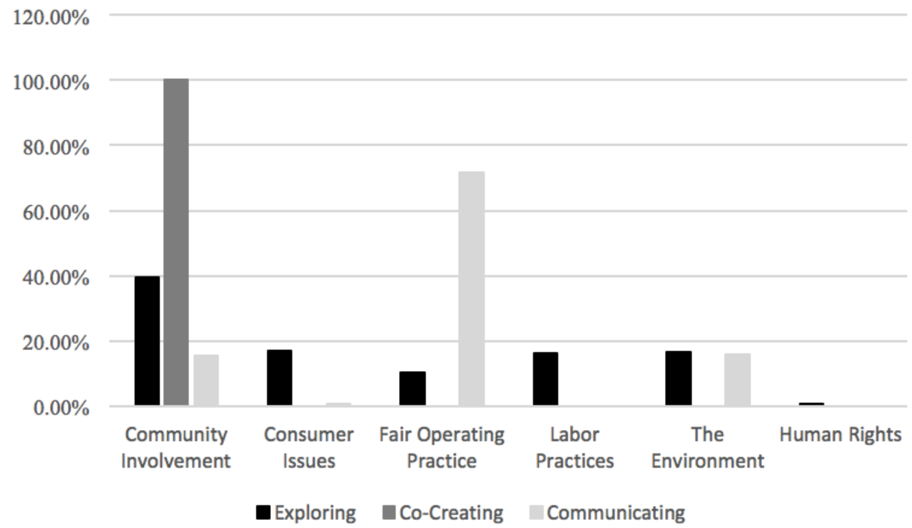

Figure 5.3 Comparison of Stakeholder Know-What

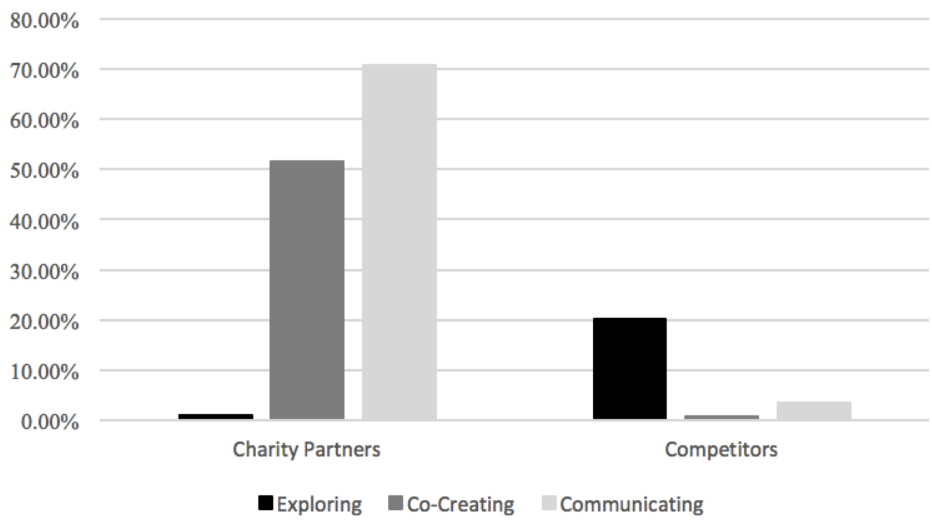

Figure 5.4 Comparison of Stakeholder Know-Who

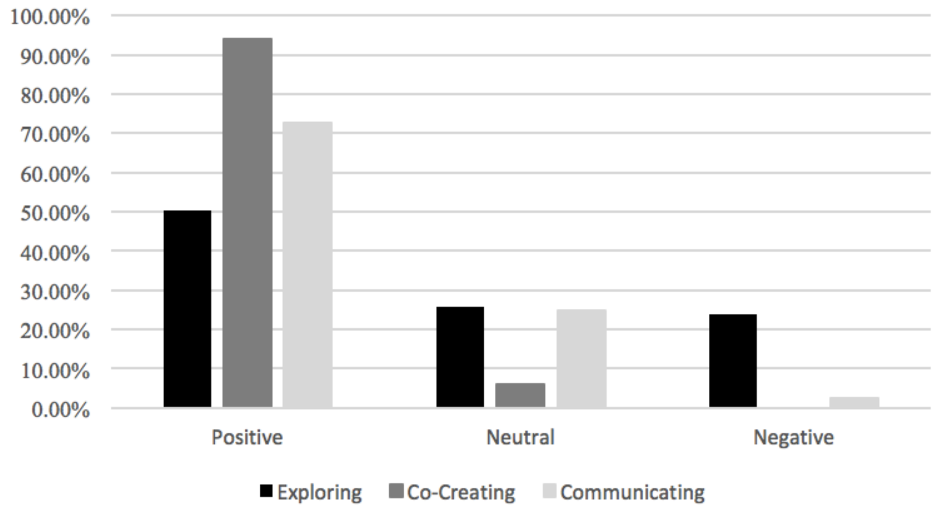

Figure 5.5 Comparison of Stakeholder Know-How (Sentiment Analysis)

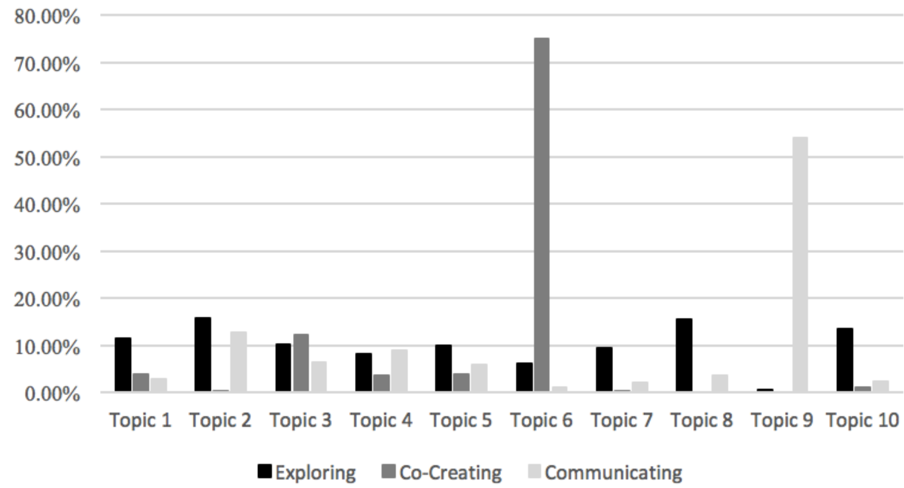

Figure 5.6 Comparison of Stakeholder Know-How (LDA Topic Modelling)

Table 5.5 LDA Results Based on 10 Topics

Topic	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
Tagged Topic	Information	Information	Action Encouraging	New service	Local Store & Group	Competition & Campaign	Store	Product	Competition & Campaign	Product
Extracted Terms	share formal hope shame	expose morning meet drink	join thankyou sweettooth task	launch shopper platform online	people pub tonnes affect	shortlist workshop shakespeare ayrshire	stores love fantastic children	vegan treat football xmas	win follow £100 instruction	bacon suffolk dutch suggest

5.6.2.2 Decision Tree Modelling Results

The results from decision tree modelling demonstrate a high feasibility of using stakeholder knowledge to classify CSI ideas. Several critical rules regarding stakeholder know-why are also uncovered (see Appendix). In the first experiment, we tested the performance of our models in different proportions of the training/test sample on the basis of 10 topics. Figure 5.7 offers the resulting model performance. Overall, our models demonstrate good performance in classifying non-ideas and improvement ideas, yet the model performance is relatively poor in predicting new ideas when the training sample is less than 20%. The best model performance in predicting new ideas is found at the ratio in which the training sample is 80% (Precision: 0.62; Recall: 0.52; F-measure: 0.57). Notably, when the ratio of the training sample is higher than 30%, the resulting models show similar performance.

In the second experiment, we tested the number of LDA topics influencing the decision tree modelling. Based on the proportion of 80% of the training sample, we ran LDA topic validation. As shown in Figure 5.8, without using any topics as predictor variables, the decision tree model demonstrates high performance in classifying non-ideas and improvement ideas. Yet, the model fails to classify new ideas. When the 10 topics were applied, the model improves the performance in classifying new ideas (Precision: 0.68; Recall: 0.48; F-measure: 0.57). At 20 topics, the model performance improves to a similar level as the 50-topic model, which is the best model of predicting new ideas (Precision: 0.81; Recall: 0.81; F-measure: 0.81).

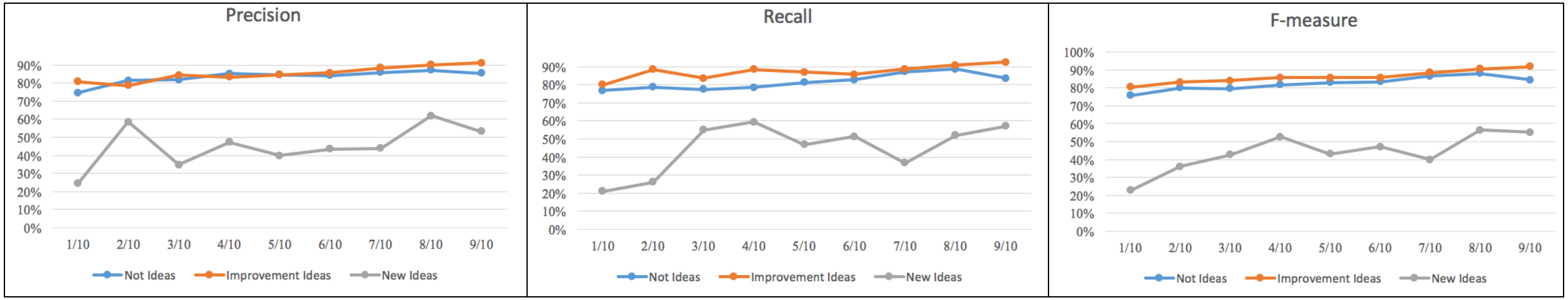

Figure 5.7 Performance Evaluation based on Variant Training Sample Size

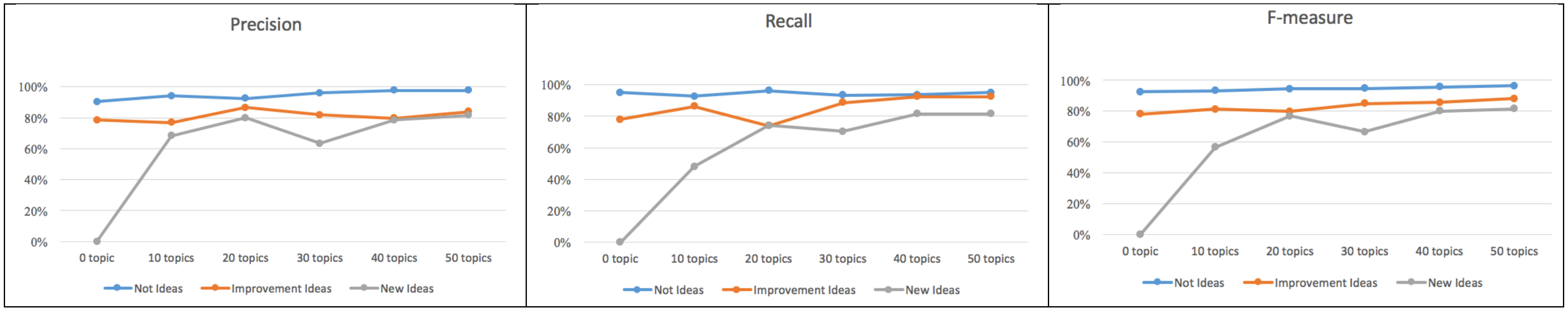

Figure 5.8 Performance Evaluation based on LDA Topic Number

5.7 Discussion

5.7.1 Theoretical Implications

This research adds to the understanding of data-driven CSI by proposing and validating a strategic framework. The framework draws on a value co-creation perspective (Prahalad and Ramaswamy 2004 ;Lusch and Nambisan 2015), highlighting that through cooperation and interaction with stakeholders, the cognitive distance regarding CSR can be reduced. The results from the literature review build linkages between the framework and prior research in terms of co-creation CSI, company value propositions of CSR (business context, institutional elements, and CSR focus areas), and stakeholder value (know-what, know-how, know-why, and know-who). Importantly, the results also indicate the literature gaps in the dimension of co-creation CSI and specific stakeholder knowledge. To bridge the gaps, a case study was further conducted and a data-driven approach exploring stakeholder value to drive co-creation CSI was demonstrated.

Our work highlights that big data captures the information regarding the company-to-stakeholder interactions and thus enables companies to identify the cognitive distance and drive co-creation CSI. In the case study, we show the use of text mining of social media dialogue data to test the feasibility of the framework in the real-life setting. Dialogue process is crucial to the next practice in value creation (Prahalad and Ramaswamy 2004). The company's CSR discourses are viewed as value propositions in which the CSR activities are embedded. The real value is, on the other hand, embedded in the stakeholder discourses. The case study examines stakeholder know-what, know-who, know-how, and know-why and links the hidden knowledge with CSR practices to create new value propositions. On the basis of the results of the case study, five fundamental propositions are clarified to guide the application of data-driven CSI.

First, a company interacts with different stakeholder groups in the business environment and engages in specific social issues. CSR activities show the company's commitment to offering solutions to social ills and improving the well-being of society (Mirvis et al. 2016). Similarly, these CSR activities are perceived by stakeholders, which is conceptualised as know-what knowledge. Lundvall and Johnson (2016, p. 112) pointed out that know-what knowledge refers to facts of the world. We argue that the facts of the world are perceived and interpreted differently, and thereby conflicts often occur due to the perception distance amongst actors. Data related to stakeholder know-what can be collected based on their social media behaviours, such as CSR-related content that stakeholders frequently share, like and respond to. We formulate the first proposition as:

Proposition 1*: Data regarding stakeholder know-what indicates the stakeholder-perceived facts of social issues. The insights related to know-what knowledge uncover the perception distance between a company and its stakeholders.*

Second, a company's CSR activities show their capabilities of allocating business resources to addressing social problems and creating social impacts through acts such as philanthropy and ethical business processes. Know-how knowledge indicates a learning-by-doing process whereby companies improve their skills and capabilities with experience over time (Garud 1997). In contrast, stakeholders develop their know-how when experiencing and interacting with CSR-relevant products, services, and processes offered in the market. Stakeholder know-how knowledge is often captured in data such as electronic word of mouth and user reviews, within which their feelings and expectations are embedded. Therefore, data related to stakeholder know-how contains potential solutions of better CSR resource arrangement. We construct the second proposition as:

Proposition 2*: Data containing stakeholder know-how is key for triggering organisational learning as it indicates a stakeholder-desired way of arranging business resources. As such, this data helps companies bridge the gap between the ideal performance and current performance.*

CSR is stakeholder-centric and involves multiple parties in the business environment and within the company (Herrera 2015). Stakeholders perceive emerging CSR issues and identify the relevant parties to take responsibility for the issues. In this way, stakeholders serve as a knowledge broker, building up linkages between the company and unidentified or uncontrollable parties (e.g. competitors). Data relevant to stakeholder know-who includes information about lead users and critical stakeholder embedment in social networks. The more the company is capable of managing such data, the closer to the market its CSR strategies will be. We formulate the third proposition as:

Proposition 3*: Data containing stakeholder know-who reveals the emerging market to serve, potential partnerships of the company and competitive market intelligence.*

In Propositions 1–3, the cognitive distance is evaluated by examining data to identify the perception gap, performance gap, and market gap. However, the subelements within the three gaps need to be further examined and prioritised to help CSI decision-making. Stakeholder know-why is established on know-what, know-how, and know-who, indicating the hidden associations amongst these three types of knowledge. The capabilities of finding representative data sources, using data analytics and discovering important patterns and rules are key to extracting stakeholder know-why knowledge. We formulate the fourth proposition as:

***Proposition 4**: Stakeholder know-why helps organisations to prioritise diverse CSR demands identified in data and to support CSI decision-making. This relies on organisations' capabilities of using data analytics to identify hidden associations and embedded rules.*

Finally, the cognitive distance indicates dissatisfactory value propositions in the company's CSR practices and discloses opportunities for implementing CSI. After the company has obtained stakeholder knowledge, it needs to evaluate if the knowledge is transformable to fit well with existing cognitive schemas (Murphy et al. 2012). We used the term knowledge internalisation (e.g. Nonaka et al. 2000) to describe the process of embodying stakeholder explicit knowledge into organisation tacit knowledge to develop CSI ideas, and we formulate the fifth proposition as:

***Proposition 5**: The insights from data should be internalised into organisation knowledge by being structured into innovative ideas for new design or improvement of CSR activities.*

5.7.2 Managerial Implications

This study has important implications for managers with an interest in understanding how CSI can be promoted through a data-driven approach. Data is of limited value before being transformed into knowledge. Prior research distinguished data, information, and knowledge, stressing that data is raw numbers and facts, information is processed data, and knowledge is the information authenticated and considered to be true (e.g. Vance 1997). In other words, knowledge is actively processed in the mind of an individual through reflection, enlightenment, and learning (Alavi and Leidner 2001). Stakeholder knowledge regarding know-what, know-how, know-who, and know-why is thus representative of their values and worldviews. The proposed framework and the five propositions can serve as an operational toolkit of knowledge discovery from stakeholder-centric data. Importantly, the stakeholder knowledge in data already possesses high market relevancy and market acceptance for CSI.

Our work also delivers practical value by demonstrating how to transform the unstructured data into knowledge using text-mining techniques. Unstructured data (e.g. social media posts) consists of 95% of big data and often remains unanalysed (Spiess et al. 2014). Prior research viewed data analytics as the critical capabilities for today's organisations to obtain competitive advantages (Kiron et al. 2012). In the case study, we used both linguistic and statistical text mining to extract known and unknown knowledge, respectively. After transforming unstructured dialogue data into a structured format, we further performed decision tree modelling to reveal the critical rules of CSI ideas. This analytical method can be applied to different types of unstructured data, helping practitioners to detect emerging trends and monitor the dynamic business environment. When social issues are viewed as wicked problems in which stakeholder demands of companies' CSR practices rapidly change, the use of text mining allows companies to detect stakeholder value more efficiently.

5.8 Conclusion

This paper presents a strategic framework and examines its practical value in supporting data-driven CSI. The findings from the literature survey and case study contribute to both academic and practical CSI research. Yet, some limitations of this research should be identified for further investigation. First, the framework proposes co-creation CSI driven by the cooperation between a company and a wide range of stakeholders such as consumers, suppliers, communities, and NGOs to discover new CSR solutions. Nonetheless, in this research, stakeholders are examined as a holistic party, considering that the main purpose of this work is to integrate data and a data-driven approach into the mechanism of co-creation CSI. We acknowledge the risk that using a general group of stakeholders tends to oversimplify CSI in the real-world context, as diverse stakeholders often possess conflicting values and compete with each other for business resources. For subsequent research, we recommend that the framework should be applied to analyse the cognitive distance between a company and diverse stakeholder groups, as well as amongst different stakeholders.

Moreover, this research develops a specific data analytical method to extract stakeholder knowledge and identify the cognitive distance within social media dialogues. Although the models built using text mining and data mining are purely data-driven and affected by the change of data sources, the mindset and analytical method are applicable to other contexts. We suggest that further research can replicate this methodology in research settings such as service innovation or new product development scenarios, not constrained to CSI, to improve the understanding of data-driven innovation.

5.9 References

Abaza, W. (2017), "Comparing Internal and External Impacts of Sustainable Innovation: An Exploratory Study," *Journal of Global Entrepreneurship Research*, 7(3), In Press.

Abrahams, A.S., Jiao, J., Fan, W., Wang, G.A. and Zhang, Z. (2013), "What's Buzzing in the Blizzard of Buzz? Automotive Component Isolation in Social Media Postings," *Decision Support Systems*, 55(4), 871–882.

Aggarwal, C.C. and Zhai, C.X. (2013), "An Introduction to Text Mining," In C.C., Aggarwal and C.X. Zhai, eds, *Mining Text Data*, US: Springer, 1–10.

Alavi, M. and Leidner, D.E. (2001), "Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues," *MIS Quarterly*, 25(1), 107–136.

AlSumait, L., Wang, P., Domeniconi, C. and Daniel, B. (2010), "Embedding Semantics in LDA Topic Models," In M. J. A. Berry and J. Kogan, eds., *Text mining: Applications and Theory*. Oxford : John Wiley & Sons, 183–204.

Altuna, N., Contri, A.M., Dell'Era, C., Frattini, F. and Maccarrone, P. (2015), "Managing Social Innovation in For-Profit Organizations: The Case of Intesa Sanpaolo," *European Journal of Innovation Management*, 18(2), 258–280.

Battisti, S. (2012), "Social Innovation: The Process Development of Knowledge-Intensive Companies," *International Journal of Services, Technology and Management*, 18(3-4), 224–244.

Le Ber, M.J. and Branzei, O. (2010a), "(Re)Forming Strategic Cross-Sector Partnerships," *Business & Society*, 49(1), 140–172.

Le Ber, M.J. and Branzei, O. (2010b), "Value Frame Fusion in Cross Sector Interactions," *Journal of Business Ethics*, 94, 163–195.

Campbell, J.L. (2007), "Why would Corporations Behave in Socially Responsible Ways? An Institutional Theory of Corporate Social Responsibility," *Academy of Management Review*, 32(3), 946–967.

Carberry, E.J., Bharati, P., Levy, D.L. and Chaudhury, A. (2017), "Social Movements as Catalysts for Corporate Social Innovation: Environmental Activism and the Adoption of Green Information Systems," *Business & Society*, In Press.

Carroll, A.B. (1979), “A Three-Dimensional Conceptual Model of Corporate Performance,” *The Academy of Management Review*, 4(4), 497–505.

Carroll, A.B. (1999), “Corporate Social Responsibility: Evolution of a Definitional Construct,” *Business & Society*, 38(3), 268–295.

Castka, P. and Balzarova, M.A. (2008). “ISO 26000 and Supply Chains: On the Diffusion of the Social Responsibility Standard,” *International Journal of Production Economics*, 111(2), 274–286.

Chien, C.-F. and Chen, L.-F. (2008), “Data Mining to Improve Personnel Selection and Enhance Human Capital: A Case Study in High-Technology Industry,” *Expert Systems with Applications*, 34(1), 280–290.

Currie, W.L. and Seddon, J.J.M. (2014), “Social Innovation in Public Health: Can Mobile Technology Make a Difference?” *Information Systems Management*, 31(3), 187–199.

Dahlsrud, A. (2008), “How Corporate Social Responsibility is Defined: An Analysis of 37 Definitions,” *Corporate Social Responsibility and Environmental Management*, 15(1), 1–13.

Desouza, K. and Smith, K. (2014), “Big Data for Social Innovation,” *Stanford Social Innovation Review*, Summer, 39–43.

Duchessi, P. and Lauría, E.J.M. (2013), “Decision Tree Models for Profiling Ski Resorts’ Promotional and Advertising Strategies and the Impact on Sales,” *Expert Systems with Applications*, 40(15), 5822–5829.

Eberle, D., Berens, G. and Li, T. (2013), “The Impact of Interactive Corporate Social Responsibility Communication on Corporate Reputation,” *Journal of Business Ethics*, 118(4), 731–746.

European Commission (2013), Guide to Social Innovation. Available at: *https://ec.europa.eu/eip/ageing/library/guide-social-innovation_en*

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D. (2015), “How “Big Data” can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study,” *International Journal of Production Economics*, 165, 234–246.

Furmańska-Maruszak, A. and Sudolska, A. (2016), “Social Innovations in Companies and in Social Economy Enterprises,” *Comparative Economic Research*, 19(3), 169–191.

Garud, R. (1997), “On the Distinction Between Know-How, Know-Why, and Know-What,” *Advances in Strategic Management*, 14, 81–101.

Goes, P.B. (2014), “Big Data and IS Research,” *MIS Quarterly*, 38(3), iii–viii.

Groot, A. and Dankbaar, B. (2014), “Does Social Innovation Require Social Entrepreneurship?” *Technology Innovation Management Review*, 4(12), 17–26.

Hanke, T. and Stark, W. (2009), “Strategy Development: Conceptual Framework on Corporate Social Responsibility,” *Journal of Business Ethics*, 85(3), 507–516.

Harazin, P. and Kósi, K. (2013), “Social Challenges: Social Innovation through Social Responsibility,” *Social and Management Sciences*, 21(1), 27–38.

van der Have, R.P. and Rubalcaba, L. (2016), “Social Innovation Research: An Emerging Area of innovation Studies?” *Research Policy*, 45(9), 1923–1935.

He, W., Zha, S. and Li, L. (2013), “Social Media Competitive Analysis and Text Mining: A Case Study in the Pizza Industry,” *International Journal of Information Management*, 33, 464–472.

Helms, W.S., Oliver, C. and Web, K. (2012), “Antecedents of Settlement on a New Institutional Practice: Negotiation of the ISO 26000 Standard on Social Responsibility,” *Academy of Management Journal*, 55(5), 1120–1145.

Hendriks-Jansen, H. (1996), *Catching Ourselves in the Act : Situated Activity, Interactive Emergence, Evolution, and Human Thought*, Cambridge, MA: MIT Press.

Herrera, M.E.B. (2015), “Creating Competitive Advantage by Institutionalizing Corporate Social Innovation,” *Journal of Business Research*, 68(7), 1468–1474.

Herrera, M.E.B. (2016), “Innovation for Impact: Business Innovation for Inclusive Growth,” *Journal of Business Research*, 69(5), 1725–1730.

Holmes, S. and Moir, L. (2007), “Developing A Conceptual Framework to Identify Corporate Innovations through Engagement with Non-Profit Stakeholders,” *Corporate Governance: The International Journal of Business in Society*, 7(4), 414–422.

Holmes, S. and Smart, P. (2009), “Exploring Open Innovation Practice in Firm-Nonprofit Engagements: A Corporate Social Responsibility Perspective,” *R and D Management*, 39(4), 394–409.

Ionescu, C. and Marga, A. (2015), “About the Conceptualization of Social Innovation,” *Theoretical and Applied Economics*, XXII(3604), 53–62.

ISO 26000, (2010), Discovering ISO 26000. *International Organization for Standardization*, 1–8. Available at: www.iso.org.

Johnson, B., Lorenz, E. and Lundvall, B. (2002), "Why all this Fuss about Codified and Tacit Knowledge?" *Industrial and Corporate Change*, 11(2), 245–262.

Kiron, D., Prentice, P.K. and Ferguson, R.B. (2012), "Innovating with Analytics," *MIT Sloan Management Review*, 54(1), 1–8.

Kolchyna, O., Souza, T.T.P., Treleaven, P.C. and Aste, T. (2016), "Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination," In M. Gautam and Y. Xiang, eds., *Handbook of Sentiment Analysis in Finance*. Albury Books.

Kolk, A. and Lenfant, F. (2015), "Cross-Sector Collaboration, Institutional Gaps, and Fragility: The Role of Social Innovation Partnerships in a Conflict-Affected Region," *Journal of Public Policy & Marketing*, 34(2), 287–303.

Korschun, D. and Du, S. (2013), "How Virtual Corporate Social Responsibility Dialogs Generate Value: A Framework and Propositions," *Journal of Business Research*, 66(9), 1494–1504.

Kotler, P. and Lee, N. (2005), *Corporate Social Responsibility: Doing the Most Good for Your Company and Your Cause*. Hoboken, NJ: John Wiley & Sons.

Kuehl, N., Scheurenbrand, J. and Satzger, G. (2016), "Needmining: Identifying Micro Blog Data Containing Customer Needs," In *Twenty-Fourth European Conference on Information Systems (ECIS)*. Istanbul, Turkey.

Kunz, W., Aksoy, L., Bart, Y., Heinonen, K., Kabadayi, S., Ordenes, F. V., Sigala, M., Diaz, D. and Theodoulidis, B. (2017), "Customer Engagement in A Big Data World," *Journal of Services Marketing*, 31(2), 161–171.

Lai, C.-S., Chiu, C.-J., Yang, C.-F. and Pai, D.-C. (2010), "The Effects of Corporate Social Responsibility on Brand Performance: The Mediating Effect of Industrial Brand Equity and Corporate Reputation," *Journal of Business Ethics*, 95(3), 457–469.

Linna, P. (2012), "Base of the Pyramid (BOP) as a Source of Innovation: Experiences of Companies in the Kenyan Mobile Sector," *International Journal of Technology Management & Sustainable Development*, 11(2), 113–137.

Lo, S.L., Cornforth, D. and Chion, R. (2015), "Identifying the High-Value Social Audience from Twitter through Text-Mining Methods," In *Proceedings of the*

18th Asia Pacific Symposium on Intelligent and Evolutionary Systems. Springer International Publishing, 325–339.

Lundvall, B.-Å. and Johnson, B. (2016), “The Learning Economy,” In B.-Å. Lundvall, eds., *The Learning Economy and The Economics of Hope*. London: Anthem Press, 107-130.

Luo, X. and Bhattacharya, C.B. (2006), “Corporate Social Responsibility, Customer Satisfaction, and Market Value,” *Journal of Marketing*, 70(4), 1–18.

Luo, X. and Du, S. (2015), “Exploring the Relationship between Corporate Social Responsibility and Firm Innovation,” *Marketing Letters*, 26(4), 703–714.

Lusch, R.F. and Nambisan, S. (2015), “Service Innovation: A Service-Dominant Logic Perspective,” *MIS Quarterly*, 39(1), 155–176.

Maiolini, R., Marra, A., Baldassarri, C. and Carlei, V. (2016), “Digital Technologies for Social Innovation: An Empirical Recognition on the New Enablers,” *Journal of Technology Management & Innovation*, 11(4), 22–28.

Manning, S. and Roessler, D. (2014), “The Formation of Cross-Sector Development Partnerships: How Bridging Agents Shape Project Agendas and Longer-Term Alliances,” *Journal of Business Ethics*, 123(3), 527–547.

Martinez, F., O’Sullivan, P., Smith, M. and Esposito, M. (2017), “Perspectives on the Role of Business In Social Innovation,” *Journal of Management Development*, 36(5), 681–695.

Mattera, M. and Baena, V. (2015), “The Key to Carving out a High Corporate Reputation Based on Innovation: Corporate Social Responsibility,” *Social Responsibility Journal*, 11(2), 221–241.

McAfee, A. and Brynjolfsson, E. (2012), “Big Data. The Management Revolution,” *Harvard Buiness Review*, 90(10), 61–68.

Mcelroy, M.W. (2003), *The New Knowledge Management: Complexity, Learning, and Sustainable Innovation*, Routledge.

Meijer, K., Frasincar, F. and Hogenboom, F. (2014), “A Semantic Approach for Extracting Domain Taxonomies from Text,” *Decision Support Systems*, 62, 78–93.

Mirvis, P., Herrera, M.E.B., Googins, B. and Albareda, L. (2016), “Corporate Social Innovation: How Firms Learn to Innovate for the Greater Good,” *Journal of Business Research*. 69(11), 5014–5021

Mulgan, G., Tucker, S., Ali, R. and Sanders, B. (2007), *Social Innovation: What it Is, Why it Matters and How it Can Be Accelerated*, Skoll Centre for Social Entrepreneurship. Available at: http://eureka.sbs.ox.ac.uk/761/.

Murphy, M., Perrot, F. and Rivera-Santos, M. (2012), "New Perspectives on Learning and Innovation in Cross-Sector Collaborations," *Journal of Business Research*, 65(12), 1700–1709.

Muthuri, J.N., Moon, J. and Idemudia, U. (2012), "Corporate Innovation and Sustainable Community Development in Developing Countries," *Business & Society*, 51(3), 355–381.

Ngai, E.W.T., Xiu, L. and Chau, D.C.K. (2009), "Application of Data Mining Techniques in Customer Relationship Management: A Literature Review and Classification," *Expert Systems with Applications*, 36(2, Part 2),.2592–2602.

Nonaka, I., Toyama, R. and Konno, N. (2000), "SECI, Ba and Leadership: A Unified Model of Dynamic Knowledge Creation," *Long Range Planning*, 33(1), 5–34.

Nooteboom, B. (2009), *A Cognitive Theory of the Firm: Learning, Governance and Dynamic Capabilities.* Cheltenham: Edward Elgar Publishing Limited.

Olejniczuk-Merta, A. (2015), "Social Innovations in the Process of System Products Development," *Journal of Economics & Management*, 22(4), 5–15.

Ordenes, F. V., Theodoulidis, B., Burton, J., Gruber, T. and Zaki, M. (2014). "Analyzing Customer Experience Feedback Using Text Mining: A Linguistics-Based Approach," *Journal of Service Research*, 17(3), 278–295.

Parmar, R., Mackenzie, I., Cohn, D., and Gann, D. (2014), "The New Patterns of Innovation," *Harvard Business Review*, 92(1–2), 86–97

Prahalad, C.K. and Ramaswamy, V. (2004), "Co-creation Experiences: The Next Practice in Value Creation. *Journal of Interactive Marketing*, 18(3), 5–14.

Pressentin, M. (2017), "Rethinking Human Resources for Social Innovation in Business: Trends - Leadership - Action Plan," *Amity Global Business Review*, 12, 52–73.

Risso, M. (2012), "Exploring Partnerships for social Innovation," *Symphonya Emerging Issues in Management*, (2), 26–36.

Roberts, D.L. and Piller, F.T. (2016), "Finding the Right Role for Social Media in Innovation Finding the Right Role for Social Media in Innovation," *MIT Sloan Management Review*, 57(3), 41–47.

Salim Saji, B. and Ellingstad, P. (2016), “Social Innovation Model for Business Performance and Innovation,” *International Journal of Productivity and Performance Management*, 65(2), 256–274.

Saul, J. (2011), “Book Highlight- Corporate Social Innovation,” *Global Business & Organizational Excellence*, 30(5), 78–87.

Schweitzer, F., Rau, C., Gassmann, O.and van den Hende, E. (2015), “Technologically Reflective Individuals as Enablers of Social Innovation,” *Journal of Product Innovation Management*, 32(6), 847–860.

Segarra-Oña, M., Peiró-Signes, A., Albors-Garrigós, J. and de Miguel-Molina, B., (2017), “Testing the Social Innovation Construct: An Empirical Approach to Align Socially Oriented Objectives, Stakeholder Engagement, and Environmental Sustainability,” *Corporate Social Responsibility and Environmental Management*, 24(1), 15–27.

Seitanidi, M.M. (2008), “Adaptive Responsibilities: Nonlinear Interactions in Cross Sector Social Partnerships,” In *Emergence: Complexity and Organization*. 51–64.

Selsky, J.W. and Parker, B. (2010), “Platforms for Cross-Sector Social Partnerships: Prospective Sensemaking Devices for Social Benefit,” *Journal of Business Ethics*, 94, 21–37.

Sen, S., Bhattacharya, C.B. and Korschun, D. (2006), “The Role of Corporate Social Responsibility in Strengthening Multiple Stakeholder Relationships: A Field Experiment,” *Journal of the Academy of Marketing Science*, 34(2), 158–166.

Sokolova, M. and Lapalme, G. (2009), “A Systematic Analysis of Performance Measures for Classification Tasks,” *Information Processing and Management*, 45(4), 427–437.

Spiess, J., T'Joens, Y., Dragnea, R., Spencer, P. and Philippart, L. (2014), “Using Big Data to Improve Customer Experience and Business Performance,” *Bell Labs Technical Journal*, 18(4), 3–17.

Thorleuchter, D., van Den Poel, D. and Prinzie, A. (2010), “Mining Ideas from Textual Information,” *Expert Systems with Applications*, 37(10), 7182–7188.

Tirunillai, S. and Tellis, G.J. (2014), “Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation,” *Journal of Marketing Research*, 51(4), 463–479.

To, C.K.M. (2016), “Collaboration Modes, Preconditions, and Contingencies in Organizational Alliance: A Comparative Assessment,” *Journal of Business Research*, 69(11), 4737–4743.

Vance, D. (1997), “Information, Knowledge and Wisdom: The Epistemic Hierarchy and Computer-based Information Systems, ” In B. Perkins and I. Vessey, eds., *The 3rd Americas Conference on Information Systems*. Indianapolis: Academic Press.

Vargo, S.L., Maglio, P.P. and Akaka, M.A. (2008), “On Value and Value Co-Creation: A Service Systems and Service Logic Perspective,” *European Management Journal*, 26(3), 145–152.

Viñals, C.R. (2013), “Introduction to Social Innovation as A New Form of Organisation in Knowledge-based Societies,” In C.R. Viñals and C.P. Rodríguez, eds., *Social Innovation:New Forms of Organisation in Knowledge–Based Societies*, NY: Routledge, 3–11.

Warnecke, T. (2017), “Social Innovation, Gender, and Technology: Bridging the Resource Gap,” *Journal of Economic Issues*, 51(2), 305–314.

Wei, X. and Croft, W.B. (2006), “LDA-Based Document Models for Ad-Hoc Retrieval,” In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Rretrieval.*

Whelan, G., Moon, J. and Grant, B. (2013), “Corporations and Citizenship Arenas in the Age of Social Media,” *Journal of Business Ethics*, 118(4), 777–790.

Witell,L., Gebauer, H., Jaakkola, E., Hammedi, W., Patricio, L. and Perks, H. (2016), “A Bricolage Perspective on Service Innovation,” *Journal of Business Research*, 79, 290–298

Xu, K., Liao, S.S., Li, J. and Song, Y. (2011), “Mining Comparative Opinions from Customer Reviews for Competitive Intelligence,” *Decision Support Systems*, 50(4), 743–754.

Yin, R.K. (2003), “Case Study Research. Design and Methods,” *SAGE Publications*, 26(1), 93–96.

Appendix:

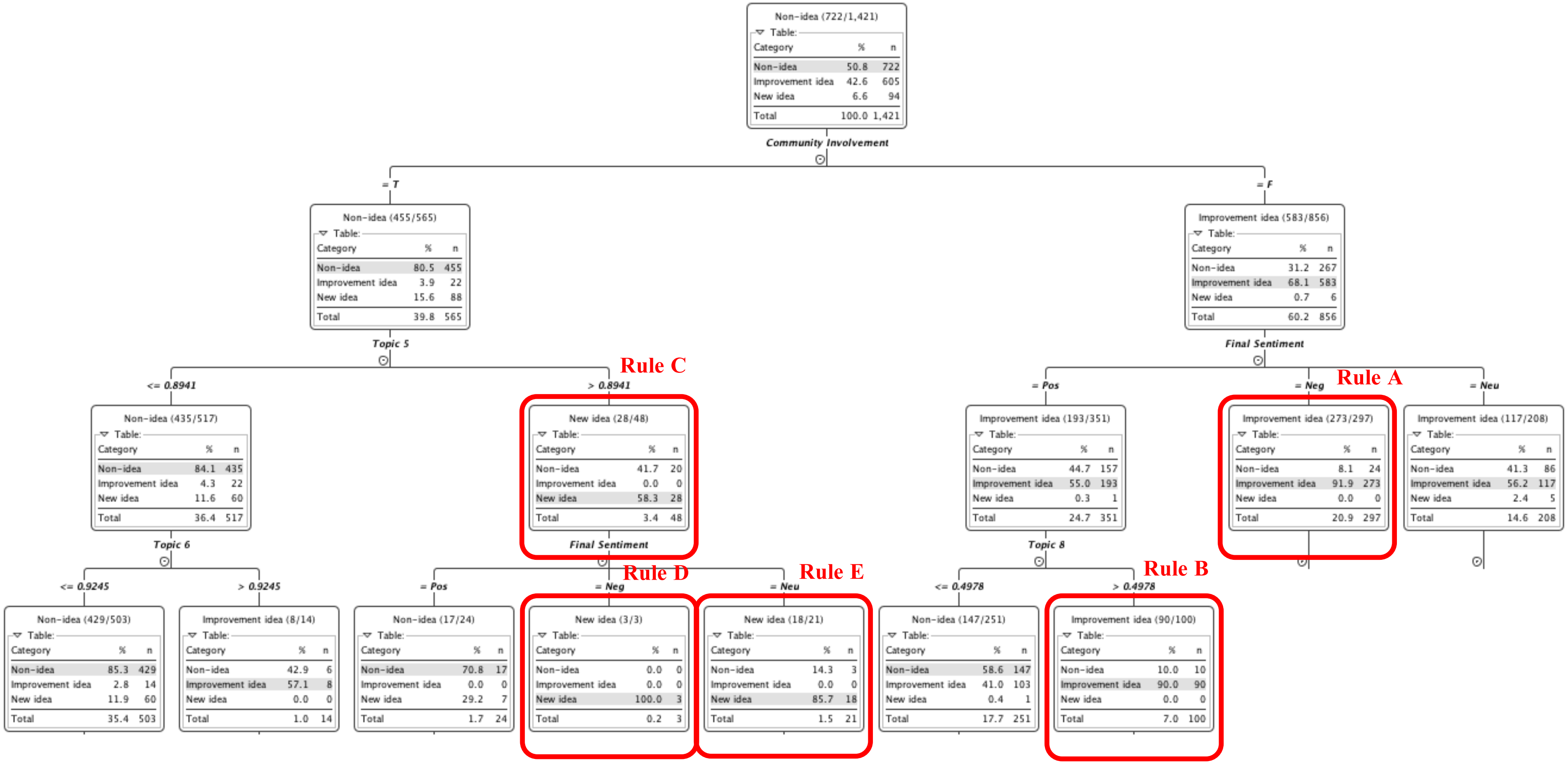

Figure 5.9 The Truncated Decision Tree Model

The Appendix shows exemplar rules regarding stakeholder know-why generated by decision tree modelling. A truncated decision tree model based on the 10-topic model is presented to demonstrate the results. The upper-level nodes are more important to tree splitting. Amongst the inputting variables, *Community Involvement* is found to be the most impactful factor. Following this are *Tweet sentiment* and *Topic 5*. We specifically focus on the nodes that predict improvement ideas and new ideas. For improvement ideas, *Tweet Sentiment* and *Topic 8* are the most relevant factors for the classification. The best prediction rules are:

Rule A: "If *Community Involvement* = False, *Tweet Sentiment* = Neg (Negative), then the probability of Improvement Ideas would be 91.9%."

Rule B: "If *Community Involvement* = False, *Tweet Sentiment* = Pos (Positive), and Topic 8 > 0.4978, then the probability of Improvement Ideas would be 90.0%."

As for the rules of new ideas, three rules are observed in Figure 5.9:

Rule C: "If *Community Involvement* = True, and Topic 5 > 0.8941, then the probability of New Ideas would be 58.3%."

Rule D: "If *Community Involvement* = True, Topic 5 > 0.8941, and *Tweet Sentiment* = Neg (Negative), then the probability of New i=Ideas would be 100.0%."

Rule E: "If *Community Involvement* = True, Topic 5 > 0.8941, and *Tweet Sentiment* = Neu (Neutral), then the probability of New Ideas would be 85.7%."

Chapter 6 Discussion and Conclusion

The final chapter is divided into two main sections. Section 6.1 focuses on the evaluation of the data-driven service systems (DDSS) framework to provide important feedback and improve understanding of the research questions. **Section 6.2** provides a summary of the thesis, discussing the research contributions, limitations and directions for further works.

6.1 Evaluation of the DDSS Framework

The development of the DDSS framework followed design science research (DSR). As stated in the research methodology section (see **Section 2.1**), the design includes building-and-evaluation iterations (Hevner 2007; March and Smith 1995). Design artefacts, such as constructs, models, methods, or frameworks, need to be evaluated with an appropriate method depending on the research objectives. Hevner (2007) specified several approaches frequently used to evaluate design artefacts, including observational, analytical, experimental, testing, and descriptive methods. Often, multiple methods are used by researchers when a design artefact contains a wide range of tasks. For instance, Osterwalder (2004) evaluated the proposed business model ontology by six methods, including comparing with the business model literature, interviewing practitioners, case studies, consulting research communities, and two field tests.

To evaluate the DDSS framework, investigation into the use of data in promoting value co-creation (**Section 6.1.1**), improvement in existing social media analytics (**Section 6.1.2**), and applicability of the framework in the field settings (**Section 6.1.3**) have been conducted. Based on Hevner’s (2007) three-cycle view of DSR, the outputs of a design artefact (design cycle) need to improve the research environment where people, organisations, and technologies are present (relevance cycle) and also contribute to the knowledge base of the design (rigour cycle). Figure 6.1 illustrates how the evaluations of the DDSS framework offer answers to the research questions outlined in **Chapter 1** and link to the knowledge base discussed in **Chapter 2**. Specifically, this chapter evaluates the utility of the DDSS framework in addressing the field problems highlighted in **Chapters 3 – 5.**

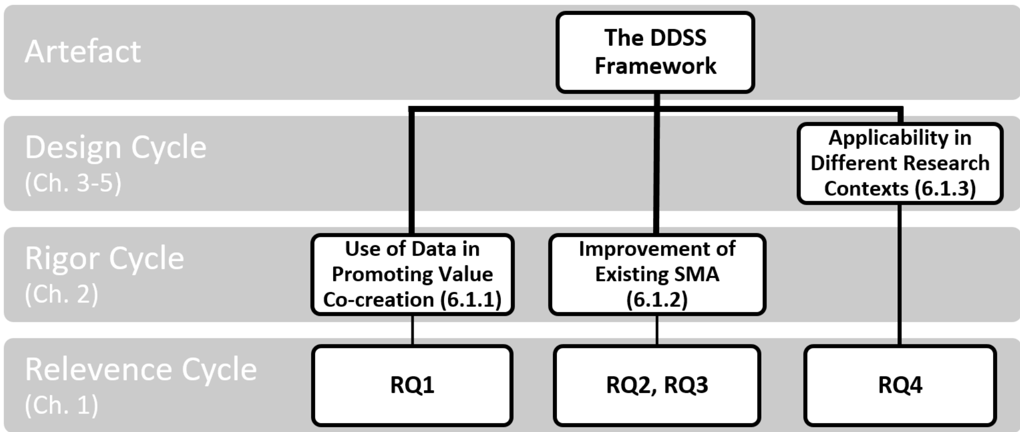

Figure 6.1 Evaluation Methods of the DDSS Framework

6.1.1 Use of Data in Promoting Value Co-creation

This research argues that big data is highly relevant for understanding value co-creation within service systems, yet it requires a high level of theorisation where the "data-driven" perspective is integrated into value co-creation. This thesis improves upon the poor definition of data-driven service systems (DDSS), highlighting how DDSS introduces changes to the system and improves the well-being of service systems through the use of big data. According to Hevner (2007), the outputs of a design artefact (design cycle) contribute to the knowledge base in both theoretical foundations and methodologies (rigour cycle). The theoretical foundations justify the theories, frameworks, constructs, models, and instantiations relevant to the artefact. In this section, the DDSS framework is evaluated through comparison with other frameworks related to big data in marketing research. In this way, the DDSS framework can be assessed based on its completeness and rigorousness in relation to the theoretical foundations.

Seven published papers in the marketing discipline were selected for the comparison. The criteria for paper selection were based on papers having topic relevancy, a well-developed theory (or a framework, a model, and a set of constructs), and operational constructs. Such comparison showed the uniqueness of the DDSS framework in covering the relevant issues through a rigorous approach suggested in Osterwalder (2004). The selected papers were published in the following journals: *Marketing Science* (1), *Journal of Marketing* (1), *Journal of Advertising* (1), *Journal of Retailing* (1), *Journal of Services Marketing* (2), and *Journal of Business Research* (1). An overview of the papers is provided in Table 6.1. For each paper, the research purpose, framework and constructs, findings and contributions are discussed.

Table 6.1 Overview of the Big-Data Frameworks in the Selected Papers

Author	Journal	Overview of the Paper
Bradlow et al. (2017)	*Journal of Retailing*	This paper offers insights into the opportunities arising from big data in the retailing sector and also highlights the importance of theory in guiding the data-driven approach for addressing retailing problems. A five-dimension retailing data model is proposed pertaining to customers, products, time, location, and channel. Also, an in-depth discussion about data analytical techniques, such as Bayesian analysis, predictive analytics, and a field experiment, is given to indicate the utility of big data (analytics) in the retailing context.
Erevelles et al. (2016)	*Journal of Business Research*	This paper examines the issues of marketing transformation driven by big data revolutions. New approaches to understanding consumer behaviour and marketing strategies through the use of customer analytics are discussed. Grounded in resource-based theory, a conceptual framework is proposed. The framework focuses on three types of company resources: physical, human and organisational. These company resources are moderated by certain elements, including (1) customer activities in the big data context, (2) customer insights capture and utilisation to enhance the company's dynamic/adaptive capabilities, and (3) value creation and sustainable competitive advantage gaining. The framework provides a better understanding of the impact of big data in marketing activities.
Kumar et al. (2013)	*Journal of Services Marketing*	This paper clarifies data-driven services marketing and discloses opportunities for improving service performance through the use of big data. A framework is provided to distinguish different types of data, including transactional data, customer data, demographic data, firm data, and attitude data. Moreover, the connections between data and services marketing metrics such as the customer-level metric (e.g. customer lifetime value, customer engagement, share of wallet, customer churn metrics) and company-level metric (e.g. service quality, human resources, operational metrics) are revealed. The proposed framework enables practitioners to make data-driven decisions and invest in the right managerial decision support systems.
Kunz et al. (2017)	*Journal of Services Marketing*	This research improves the understanding of customer engagement in the big data context. Drawn on the value co-creation perceptive, a strategic framework is offered that integrates

		the customer perspective to the company's offerings. The framework entails the important elements involved during service interactions: firm resources, data, process, timeline, and goals for engagement. Moreover, the dynamic factors are specified, such as customers' motives, situational factors, and preferred engagement styles. The findings from this paper shed light on data-driven customer engagement and provide practical value for managing the dynamic and interactive value co-creation processes.
Malthouse and Li (2017)	*Journal of Advertising*	The authors introduce how big data shapes and creates new opportunities in advertising research. A framework is provided to uncover the big data generated from consumer behaviour (e.g. dialogue behaviour, shopping behaviour, and use behaviour), brand actions (e.g. ad message, webcare) and brand outcomes (e.g. purchase, loyalty, brand equity) and to clarify the associations amongst these three dimensions. Several research questions are highlighted to give directions to further research, such as developing and testing theories, identifying insights, and optimising the delivery of messages.
Wedel and Kannan (2016)	*Journal of Marketing*	The authors examine the potential value of marketing analytics to support marketing decisions. A framework is provided covering the data sources and the marketing analytics for supporting decision-making regarding CRM, marketing mix, personalisation, customer privacy, and data security. This paper offers suggestions for a data-driven approach to marketing practice and points out the importance of data analytics as a discipline in marketing education.
Rust and Huang (2014)	*Marketing Science*	This paper discusses marketing transformation driven by the improvement of information technology. A conceptual framework is proposed, explaining how the advance in IT leads to service revolutions through ubiquitous customer communication and big customer data. The framework also uncovers the critical management issues, such as dynamic customer interactions, customer relationship management over time, and use of marketing analytics in reacting to service revolutions.

Table 6.2 details the comparison amongst the prior works based on the management blocks of the DDSS framework: *service systems*, *data environment*, and *decision-making*. Each element was evaluated according to its importance and relevance to the selected papers, with a scale from 0 to 3. The higher score an element obtains, the higher importance the element holds. For example, all three management blocks are covered and modelled in Kunz et al.'s (2017) customer engagement framework using big data, with a score of 3 given to each element. In contrast, Kumar et al.'s (2013) framework also covers the three elements, but the element *service systems* stays noncommittal on the description, with a score of 1 given. The findings in Table 6.2 indicate that, perhaps not surprisingly, the elements *data environment* and *decision-making* are mentioned in all selected papers. As for the element *service systems*, three out of seven papers mentioned (or even modelled) it in constructing their frameworks.

Table 6.2 The DDSS Management Blocks Compared to the Prior Research

Authors	Service Systems	Data Environment	Decision-Making
Bradlow et al. (2017)	0	3	1
Erevelles et al. (2016)	0	3	1
Kumar et al. (2013)	1	3	2
Kunz et al. (2017)	3	3	3
Malthouse and Li (2017)	0	3	1
Wedel and Kannan (2016)	0	3	3
Rust and Huang (2014)	3	3	3
Number of Times the Component is Mentioned	3/7	7/7	7/7

* Criteria: 0= Element not existing, 1= Element mentioned, 2= Element discussed, 3= Element modeled

Table 6.3 shows how the nine operational components of the DDSS framework (*actors*, *resources*, *activities*, *context*, *time*, *big data*, *BDA*, *value*, and *decision metrics*) are covered in the selected papers and how they are relevant to these works. As shown, all operational components are mentioned in more than half of the selected papers, except for the component *context*, which is only mentioned in three papers. The components *big data*, *BDA*, and *value* are present in all the papers. *Value* is modelled in most of the papers (five out of seven), but only mentioned in two papers. Notably, each paper takes a different aspect of value. For instance, Kumar et al. (2013) and Kunz et al. (2017) examined the value regarding customer lifetime value and customer engagement value. Wedel and Kannan (2016) proposed two types of value in their work: information value and decision value.

The components *resources* and *activities* are less evident constructs in the selected papers, with two papers modelling *resources* as a construct and only one paper modelling *activities* into the framework. However, each paper focuses on specific sub-dimensions of business resources and activities rather than using a generic category of *resources* and *activities*. For example, Bradlow et al.'s (2017) retailing data dimensions contain components such as channel and product, and these two can be categorised as a resource in the DDSS framework.

The results of the comparison with prior research show that the DDSS framework provides a high level of abstraction and accommodates the constructs identified in the previous big data research in marketing. Therefore, the proposed framework manifests its usefulness in improving the theoretical foundations.

Table 6.3 The DDSS Operational Components Compared to the Prior Research

Authors	Actors	Resources	Activities	Time	Context	Big Data	BDA	Value	Decision Metrics
Bradlow et al. (2017)	3	1	1	3	0	3	3	1	3
Erevelles et al. (2016)	3	3	3	0	0	3	3	3	0
Kumar et al. (2013)	3	1	1	3	0	3	3	3	3
Kunz et al. (2017)	3	3	2	3	3	3	3	3	3
Malthouse and Li (2017)	3	0	0	3	3	3	3	1	3
Wedel and Kannan (2016)	0	0	0	1	0	3	3	3	3
Rust and Huang (2014)	3	0	0	3	2	3	3	3	2
Number of Times the Component is Mentioned	6/7	4/7	4/7	6/7	3/7	7/7	7/7	7/7	6/7

* Criteria: 0= Element not existing, 1= Element mentioned, 2= Element discussed, 3= Element modeled

6.1.2 Improvement of Existing Social Media Analytics

In **Section 6.1.1**, the DDSS framework is evaluated based on the theoretical foundations. Hevner (2007) suggested that the new design also serves as an extension of prior knowledge with respect to methodologies (e.g. data analyses, measures, validation criteria). In this section, the DDSS framework is applied to examine the current BDA as a way to evaluate its usefulness, relevance, and extensibility (rigour cycle). The evaluation is especially centred on social media analytics (SMA), as this thesis is conducted to examine social media dialogue data for the understanding of DDSS.

Twelve SMA vendors have been selected to validate the DDSS framework, and their details are provided in Table 6.4. These vendors have been highlighted as the most significant SMA providers in the recent Forrester report, in which the vendors were assessed based on a 30-criteria measure (Samantha and Pilecki 2016). The selection criteria of vendors are given as follows: (1) the vendor has revenues of at least \$15 million generated from their social media analytics product, (2) the vendor has at least 50 enterprise clients whose annual revenue is at least \$1 billion; and (3) the vendor's product offers the functions of data collection, data analysis, actionable insights, producing and presenting (Samantha and Pilecki 2016).

To commence the analysis, an investigation of the SMA vendors' product reports and official sites was conducted to understand their big data capabilities. This was done following Fan and Gordon's (2014) CUP framework: capture, understand and present social media data. Table 6.4 shows the results of the SMA vendor analysis and provides information in terms of data sources (capture), social media analytics functionality (understand), and marketing solutions (present). As shown, a high similarity is found amongst the vendors' key abilities. All the selected SMA products are capable of processing multiple data sources, including social media, online review sites, or news sites. To deal with the various data sources, all the SMA products are capable of investigating both structured and unstructured data. The built-in algorithms for unstructured data analysis include text analytics, image analytics, and sentiment analysis; and for processing structured data, the algorithms include user profiling, user demographics and geo-location analysis, and social metrics.

On the basis of the SMA product information, an in-depth examination was implemented to relate the SMA features to the DDSS framework (see Figure 6.2). In the upper part of Figure 6.2, a cross-tabulation is constructed using the operational components of *service systems* (rows) and *data environment* (columns). The cross-tabulation depicts what types of data and SMA are useful for understanding the *actors*, *resources*, *activities*, *time*, and *context* of social media. For instance, an actor's experienced resources and activities with a company are embedded in the unstructured data, such as texts, pictures, and videos. These types of data require SMA techniques such as text analytics, image analytics or the SMA vendor's domain-specific mathematical models (e.g. social reputation scoring, crisis alert systems) to extract hidden information.

The upper cross-tabulation analysis is subsequently associated with the lower part of *decision-making*, where actor value and correspondent decision metrics are specified. For example, an actor's online personal details are captured in the data of user profiles (e.g. age, gender, nationality), and SMA techniques such as user demographics analysis can be used to examine the data. Insights from the SMA help enhance company value in terms of brand awareness gained from more accurate social media advertisement or contents, and the decision metrics used for evaluating the company value include social performance metrics (e.g. number of impressions and number of fans/members).

As stressed earlier, big data is generated through actor interactions within service systems. Social media data records the history of user interactions between a company and its customers (or stakeholders), and between customers and other social media users. However, most of the SMA products are designed to capture, analyse and present customer-centric data (e.g. customer posts, likes, shares) but ignore the company-generated data. Amongst the examined SMA products, only three out of twelve (Clarabridge, Salesforce, Sysomos) highlight the function of analysing companies' data, yet the analysis is limited to structured data (e.g. service agents' response times). An in-depth investigation on how company-generated data is associated with specific customer outcomes is absent. Moreover, the examination of the SMA products also reveals the lack of a mutual-beneficiary perspective on decision-making. As shown in the decision-making block, the SMA products focus heavily on the value as well as decision metrics tailored to understand company benefits such as campaign performance and marketing ROI. How customer value can be improved and evaluated is only examined in the aspect of customer satisfaction.

Based on the above analysis, it is suggested that the DDSS framework can be used for SMA design and improvement by clarifying linkages among *service systems*, *data environment*, and *decision-making*. Also, the framework adds to the analytics-based models (e.g. Fan and Gordon 2014; Stieglitz and Dang-Xuan 2013) discussed in **Chapter 2** by providing a holistic view of SMA which allows practitioners to investigate the connections between user behaviour and data, between data and data analytics, and between analytics-enabling insights and business strategies. More specifically, the DDSS framework points out not only "what" user interaction data to use, but importantly "how" to use big data and "where" the data-driven insights will provide a contribution.

The dialogue-mining approach proposed in this thesis is derived from the DDSS framework. This approach is closely related to existing SMA and has been tested by the three manuscripts. The Twitter datasets used in the three papers are representative for the information regarding actors, resources, activities, time and context in the service systems. Twitter data has been widely analysed in the business setting for capturing real-time customer insights and market trends. For example, Chapter 3, and 4 added to the complaint-handling operations via social media and examined the resources, activities, temporal and contextual factors during service recovery by extracting the information from actor dialogues. The dialogue-mining approach provides an integrated analytical pipeline facilitating text analytics, sentiment and emotion mining, and time-series analysis. This approach serves as solutions to better investigate value co-creation amongst actors through uncovering crucial insights in the interrelated messages. The insights from dialogue-mining provide concrete suggestions for decision-making and improve the understanding of service recovery and corporate social innovation, thus, improving the well-beings of service systems.

In the next section, an in-depth discussion is conducted to evaluate how the three manuscripts use the DDSS framework to address the field questions and contribute to service research.

Table 6.4 List of the Selected SMA Providers

No.	SMA Provider	SMA Product(s)	Captured Data Source	SMA Functions	Marketing Solutions
1	Brandwatch	Brandwatch Analytics, Brandwatch Audiences	Blog, Twitter, Facebook, Forums, Review, Images, Videos, News, etc.	• Image Analytics • ROI Measurement • Sentiment Analysis • Share of Voice • Social Metrics[8] • Text Analytics • Time Series Analysis • User Demographics and Geo-location • User Influence Scoring • User Profiling	• Brand Reputation Management • Competitor Benchmarking • Influencer Identification • Market Research • Market Trend • Target Audience Understanding
2	Cision	Social Software	Twitter, Facebook, Blogs, Websites, Pinterest and videos, forums, Instagram, etc.	• Sentiment Analysis • Share of Voice • Social Metrics • Text Analytics • Time Series Analysis • User Demographics and Geo-location • User Profiling	• Brand Reputation Management • Competitor Benchmarking • Influencer Identification • Lead Identification[9] • Market Trend • Social Performance

[8] **Social Metrics** often include the information about how many times a specific user-generated content is viewed, shared mentioned, interacted.
[9] **Lead Identification** is based on users' social media behaviour and social conversations to find and connect with people who express an interest in companies' products/services and their intent to purchase.

3	Clarabridge	CX Social	Facebook, Twitter, Instagram, blogs, forums, reviews, and news sites, photos, etc.	• Image Analytics • ROI Measurement • Sentiment Analysis • Share of Voice • SLA Monitoring[10] • Social Metrics • Tagging and Routing Analytics • Text Analytics • Time Series Analysis • User Demographics and Geo-location • User Profiling	• Brand Reputation Management • Crisis Management • Influencer Identification • Market Trend • Webcare Performance
4	Crimson Hexagon	Crimson Hexagon Platform	Facebook, Twitter, Instagram, blogs, forums, reviews, and news sites, photos, YouTube etc.	• Emotions and Sentiment Analysis • Image Analytics • Share of Voice • Social Metrics • Text Analytics • Time Series Analysis • User Demographics and Geo-Location • User Profiling	• Brand Reputation Management • Campaign Measurement • Competitor Benchmarking • Crisis Management • Market Trend • Influence Identification • Lead Identification • Target Audience Understanding
5	NetBase	NetBase Pro, Netbase Enterprise	Facebook, Twitter, Instagram, blogs,	• Image Analytics • Sentiment and Emotion	• Brand Reputation Management

[10] **SLA Monitoring** helps manage the webcare and conversational interactions efficiently by informing the team regarding how long mentions have been waiting for an answer, avoiding the low response rate and delayed replies. Moreover, it providers insights about the quality and quantity of engagement with tone of voice, and unique users serviced.

			forums, reviews, and news sites, photos, YouTube etc.	Analysis • Share of Voice • Social Metrics • Text Analytics • Time Series Analysis • User Demographics and Geo-Location • User profiling	• Campaign Measurement • Competitor Benchmarking • Crisis Management • Influence Identification • Target Audience Understanding
6	Networked Insights	Kairos	Twitter, Facebook, YouTube, WordPress, blogs, forums, review, new sites etc.	• Marketing Mix Econometric Modeling • Sentiment Analysis • Share of Voice • Social Metrics • Text Analytics • Time Series Analysis • User Demographics and Geo-Location • User profiling	• Brands/Product Management • Campaign Performance • Competitor Benchmarking • Market Trend • Crisis Management • Target Audience Understanding
7	Oracle	Social Cloud	Facebook, Twitter, Instagram and Weibo, blogs, consumer review sites, video sites etc.	• Dynamic Link Tracking • Sentiment Analysis • Share of Voice • Social Metrics • Text Analytics • Time Series Analysis	• Brands/Product Management • Campaign Performance, • Competitor Benchmarking • Consumer Intent • Influencer Identification • Market Trend • Target Audience Understanding
8	Prime Research	Media Insight Suite	Facebook, Twitter, Instagram, blogs, forums, reviews,	• Sentiment Analysis • Share of Voice • Social Metrics	• Brands/Product Management • Campaign Performance, • Competitor Benchmarking

			and news sites, photos etc.	• Text Analytics • Time Series Analysis	• Influencer Identification • Market Trend • Target Audience Understanding
9	Salesforce	Social Studio	Facebook, Twitter, Instagram, LinkedIn, blogs, comments, consumer review sites, video sites etc.	• Image Analytics • Sentiment Analysis • Share of Voice • Social Metrics • Social Network Analysis • Text Analytics • Time Series Analysis • User Demographics and Geo-Location • ROI Measurement	• Brands/Product Management • Campaign Performance • Competitor Benchmarking • Crisis Management • Webcare Performance • Influence Identification
10	Sprinklr	Sprinklr Experience Cloud	Facebook, Twitter, Instagram, YouTube, blogs, comments, consumer review sites etc.	• Content Engagement Metrics • Image Analytics • Sentiment Analysis • Share of Voice • Social Metrics • Text Analytics • Time Series Analysis • User Demographics and Geo-Location • User Profiling	• Brands/Product Management • Campaign Performance • Competitor Benchmarking • Crisis Management • Influencer Identification • Market Trend • Target Audience Understanding • Team Performance
11	Synthesio	Synthesio Platform	Facebook, Twitter, Instagram, YouTube, Slideshares, blogs, comments,	• ROI Measurement • Sentiment Analysis • Share of Voice • Social Metrics • Social Reputation Score	• Brand Reputation Management • Brands/Product Management • Complaint Management • Crisis Management

			consumer review sites etc.	• Text Analytics • Time Series Analysis • User Demographics and Geo-Location	• Influencer Identification • Market Research • Market Trend
12	Sysomos	Sysomos Platform	Facebook, Instagram, Twitter, YouTube, News, forums, blogs	• Sentiment Analysis • Share of Voice • SLA Monitoring • Social Metrics • Social Network Analysis • Text Analytics • Time Series Analysis • User Demographics and Geo-Location	• Brands/Product Management • Campaign Performance • Competitor Benchmarking • Crisis Management • Influencer Identification • Market Trend • Target Audience Understanding • Webcare Performance

Service Systems	Data Environment	
	Big Data	**BDA**
Actors	User Profiles User Networks Frontline Employee Action Logs	User Demographics and Geo-location User Profiling User Influence Scoring Social Network Analysis SLA Monitoring
Resources **Activities**	*Data Types:* Texts, Photos, Videos, User tags, likes, shares etc. *Data Content:* Complaints, Compliments, Purchase Intent, Queries etc.	Sentiment and Emotions Analysis Text Analytics Image Analytics Marketing Mix Econometric Modeling Content Engagement Metrics Social Reputation Score Social Metrics ROI Measurement Tagging and Routing Analytics Dynamic Link Tracking
Time	Historical and Real-Time Data	Time Series Analysis, Trend Analysis
Context	Competitors' and their Customer Data	Share of Voice

Decision-Making			
Customer Value	**Decision Metrics**	**Company Value**	**Decision Metrics**
Customer Satisfaction	Social Performance Metrics	Brand Reputation Management Brands/Product Management Crisis Management Competitor Benchmarking Consumer Intent Influencer Identification Target Audience Understanding Market Trend Detection	Marketing ROI Market Research Campaign Performance Metrics Content Performance Metrics Webcare Performance Metrics Social Performance Metrics Share of Voice Team Performance Metrics

Figure 6.2 Mapping the SMA onto the DDSS Framework

6.1.3 Applicability in Different Research Contexts

The DDSS framework has been applied to address field problems in two different research contexts (see **Chapters 3**, **4** and **5**). In this section, a thorough investigation on how the DDSS framework manifests its usefulness in the three applications is implemented. First, the discussion draws on the management blocks of the DDSS framework, consisting of *service systems*, *data environment*, and *decision-making*. The assessment of the three blocks enables researchers to achieve the following aims:

- understanding real-life actor dialogical interactions and value co-creation mechanisms in *service systems*,
- capturing dialogue data and uncovering actionable insights using data analytics in the *data environment*, and
- assessing new value propositions and corresponding actor value for *decision-making*.

The second evaluation is conducted on the nine operational components of the DDSS framework: *actors*, *resources*, *activities*, *time*, *context*, *big data*, *big data analytics* (BDA), *value*, and *decision metrics*. These components can serve as a strategic tool, a conceptual model, or an IT prototype to address the field problems.

6.1.3.1 Evaluation of the First Paper

The first paper, provided in **Chapter 3**, offers insights on the management of dynamic factors (e.g. competitor and other user engagement) during service recovery, and it demonstrates the use of text mining for extracting insights (e.g. service failures, service recovery activities and customer emotions change) from the dialogue data.

Figure 6.3 describes the DDSS management blocks of the first paper. As shown in the *service systems* block, the main goals of such systems are achieved when successful service recovery occurs and benefits both the company (e.g. avoid customer churn) and the complainers (e.g. customer satisfaction). The process whereby service recovery activities are conducted to restore customer satisfaction is interpreted as a value co-creation process (Xu et al. 2014). Identifying the dynamic factors that hinder value co-creation between the company and the complainers is the vital task of the first paper.

In the *data environment* block, the dialogue data between service agents and dissatisfied customers depict the real-time actor interactions during service recovery. Dialogue data reflects the truth occurring in *service systems* and contains crucial information about the dynamic issues of service recovery. Therefore, dialogue data is indicative of the problematic service recovery offerings (value propositions), customer emotional responses bonded with each interaction (outcomes of value co-creation), and social influence amongst social media users (dynamic factors). In this block, another mechanism of value co-creation takes place between the company and BDA. BDA serves as a value integrator that extracts insights from the interactive dialogues and facilitates the company's value co-creation (e.g. learning about weak links in the service processes). The insights generated in the *data environment* are applied to the *decision-making* block in forms of new value propositions that enable companies to offer better service recovery offerings and improved complaint handling practices.

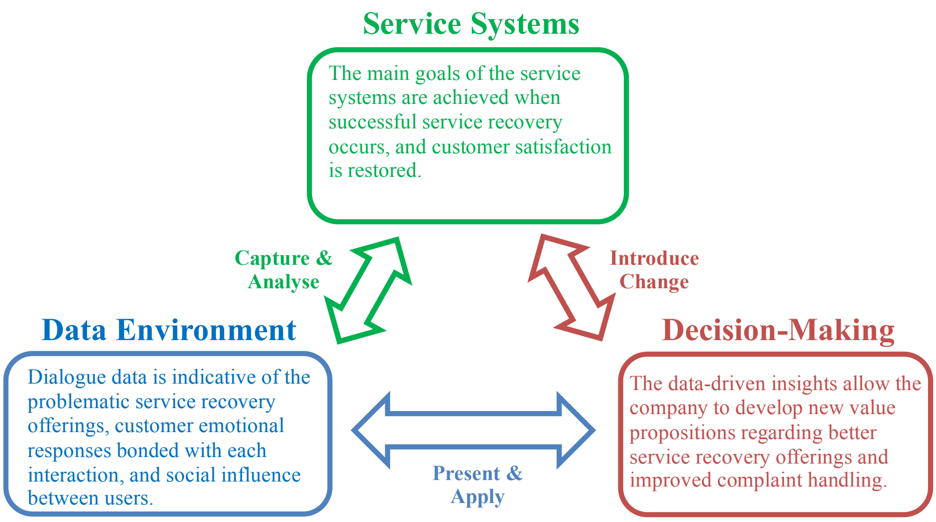

Figure 6.3 Evaluating the DDSS Management Blocks in the First Paper

The DDSS operational components serve as a strategic tool enabling researchers and practitioners to investigate the problem area. The application of the nine components in addressing social media service recovery issues is justified as follows:

- **Actors:** The actors who can influence the well-being of service systems include the company, complainers, other customers (both observers and participants), and competitors.
- **Resources**: Two types of resources are highlighted during service recovery, including resources leading to service problems and resources for problem rectification. The former includes problematic resources that cause service value failure (e.g. defective products, price, and promotion issue) and co-creation failure (e.g. customer knowledge and skills), and the latter includes service recovery resources (e.g. refund, solution).

- **Activities**: Similarly, the activities during service recovery include activities leading to service problems (e.g. delayed delivery, staff mistreatment) and activities for problem rectification (e.g. apology, channel direction).
- **Time**: The temporal dimension in service recovery can be examined based on the interaction frequency and duration between the company and the complainers.
- **Context**: The contextual factors are uncontrollable for the company but have an impact on service recovery outcomes. These factors include external drivers of service failures (e.g. weather, fraud, customers' fault) and uncontrollable actors, resources, and activities involved in service recovery processes (e.g. social influence, customer knowledge, and skills).
- **Big Data**: The dataset used was collected from a UK telecoms company's Twitter customer care during a six-month period. A total of 17,125 complaint handling dialogues were identified from 36,954 Twitter dialogues. The dataset was rich in information relevant to the research issues and representative for the service recovery scenario.
- **BDA**: An ontological approach was applied to analyse service failures and service recovery activities mentioned in the Twitter dialogues. Text mining was used to extract the embedded information (variables) following the ontology. Finally, the variables resulting from the text mining were examined by logistic regression models to reveal the impacts of dynamic factors, service failures, and service recovery strategies on customer post-failure satisfaction. The research design covers the critical stages of BDA, including data acquisition, data exploitation and data-driven insights offered.

- **Value**: The findings in the first paper indicate the main effects and moderate effects of the dynamic factors (e.g. multi-actor engagement, customer co-recovery) on the customer service recovery experience. The dialogue data contributes to the company's value creation with respect to lower customer churn (value-in-exchange), improved service practice (value-in-use) and better customer support in a networking context (value-in-context). On the other hand, customer value is enhanced regarding customer satisfaction (value-in-use) and altruism and vengeance (value-in-context).
- **Decision Metrics**: To evaluate customer satisfaction (customer value-in-use), tailored decision metrics were developed based on customer sentiment change within complaint handling dialogues. This paper also provides summary statistics of service failures (company value-in-use) mentioned in customer tweets, thus, allowing the company to take actions to improve service quality.

6.1.3.2 Evaluation of the Second Paper

The second paper, provided in **Chapter 4**, offers a dialogue-mining method, which allows examination of the linguistic and semantic, process, and relationship dimensions of dialogues in the context of social media webcare. Also, this paper addresses two important gaps in service recovery regarding process recovery and customer recovery. Figure 6.4 describes the DDSS management blocks of the second paper.

This paper shares some similarities with the first paper. Beyond the customer post-recovery satisfaction discussed in the first paper, *service systems* can be enhanced through the process recovery that the company learns from customer experience to improve service practice and obtain competitive advantages (Michel et al. 2009). Therefore, value co-creation in *service systems* functions in two ways: customer recovery and process recovery. Notably, in the process recovery, the mechanism of value co-creation is driven by customer complaints as value propositions, and the real value is achieved when the company uses complaints to review and redesign service delivery processes and improve performance.

In the *data environment* block, the dialogue data between the company and the customers are investigated. In contrast to the first paper, this paper focuses on developing a data analytical framework that examines interactive dialogues in the dimensions of linguistic and semantic (e.g. acts, episodes), process (e.g. sequences), and relationship. The dialogue-mining framework is designed to transform complicated and noisy user interaction data into a clean and interpretable structure. The framework is applied to address the issues of customer recovery and process recovery in *service systems*. The insights generated from mining dialogues contribute to the block of *decision-making* by uncovering problematic service / service recovery offerings (existing value propositions) and providing the directions for new value propositions, such as improved resource allocations amongst complaint handling channels.

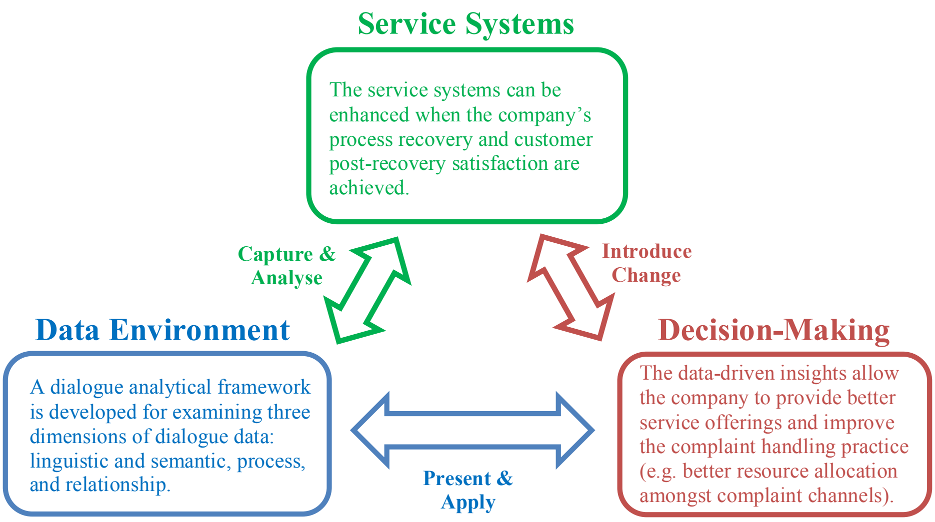

Figure 6.4 Evaluating the DDSS Management Blocks in the Second Paper

The operational components of this paper are similar to those of the first paper, yet some adjustments are made in the components *big data* and *BDA*:

- **Big Data**: The Twitter dataset used in this paper was collected from a UK grocery retailer and contains 7,201 dialogues.
- **BDA**: Analysing dialogue data using BDA was examined in prior research only to a limited extent. Most of the previous marketing research used qualitative analysis methods, yet these methods often could only process a small sample size. To advance dialogue analysis techniques, a framework is proposed and translated into a dialogue-mining apparatus that demonstrates an analytical process – dialogue data collection, information extraction, process modelling and relationship evaluation. This approach facilitates text mining and process mining to deal with a large amount of dialogue data and extracts embedded insights speedily.

6.1.3.3 Evaluation of the Third Paper

The third paper, provided in **Chapter 5**, suggests a data-driven approach to enhance companies' performance in corporate social responsibility (CSR) and drive corporate social innovation (CSI). Figure 6.5 shows the DDSS management blocks of the third paper. In the *service systems* block, the value co-creation is achieved when companies conduct CSI and allocate business resources in a way that meets stakeholders' expectations. CSR activities (value propositions) within service systems are often viewed differently by various stakeholders, which is identified as cognitive distance. It is important for companies to reduce cognitive distance and conduct CSI to achieve shared value. Therefore, engaging multiple stakeholders in CSR activities and obtaining stakeholder knowledge to co-shape new value propositions becomes an important strategy to enhance the well-being of such service systems.

In the *data environment* block, the dialogue data between the company and its stakeholders reveals the worldviews of both parties. This paper facilitates a data-driven approach to extracting the worldviews of both parties within dialogues and identifying the cognitive distance. Several text-mining techniques were applied to examine stakeholder knowledge regarding know-what, know-how, know-who, and know-why, thus enabling the company to assess the cognitive distance between itself and its stakeholders. In the *decision-making* block, stakeholder knowledge is internalised as innovative ideas that drive the company's CSR improvement or even CSI (new value propositions).

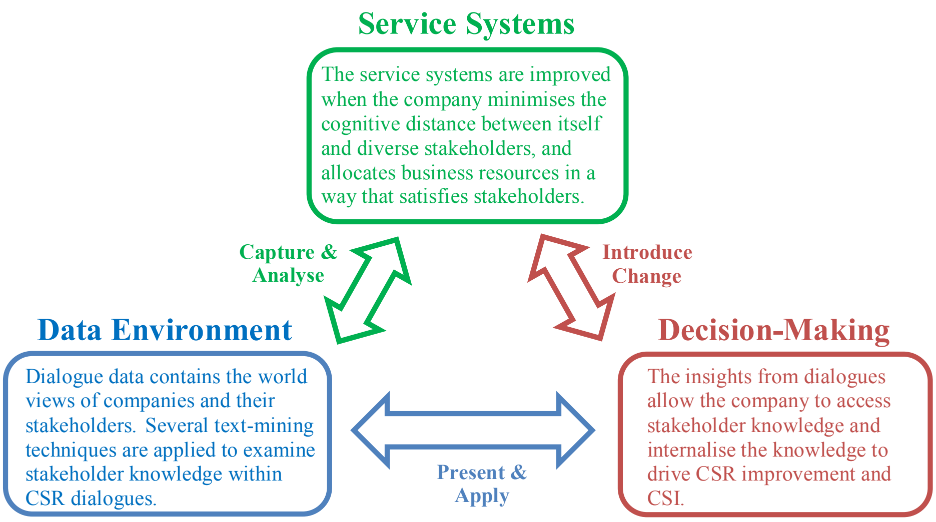

Figure 6.5 Evaluating the DDSS Management Blocks in the Third Paper

The nine operational components are justified as below:

- **Actors**: The actors who are concerned with CSR include the company, customers, local communities, suppliers, competitors, NGOs, governments and other stakeholders.
- **Resources**: The resources invested in the company's CSR are aimed to improve the well-being of stakeholders and create social impact. The CSR resources include the company's owned resources, such as eco-friendly products, and the company's facilitated resources, such as customer donations and public funds.
- **Activities**: The CSR activities can be examined through two types of activities: resource exchange, through which stakeholders obtain products, services or companies' goodwill (e.g. healthy food offering, eco-friendly store), and resource integration, including local community support.
- **Time**: The time span of a CSR issue mentioned by stakeholders indicates if the issue is a temporary or long-lasting trend.

- **Context**: The contextual elements within such service systems include the changing ethical demands from stakeholders (e.g. green policy, new definitions of healthy food) or competitors' or other organisations' actions of better CSR practice (e.g. pay rise for employees).
- **Big Data:** As with the second paper, the UK grocery retailer's Twitter dataset was used, including 7,201 dialogues.
- **BDA**: Two text-mining techniques were employed for extracting stakeholder knowledge. Linguistic-based text mining was used to identify the well-structured knowledge of stakeholders, such as CSR perceptions, CSR performance, and critical parties of specific CSR issues. Statistical text mining was used to identify the unknown knowledge mentioned in the stakeholders' tweets. The results of text mining were applied to decision tree modelling to understand the pattern of CSI ideas.
- **Value**: Mining CSR dialogues serves as an efficient method for accessing stakeholder knowledge and changing CSR trends. Value is often perceived differently and there are conflicting perspectives amongst various stakeholders. For example, customers' value-in-context may be altruism gained from donations, while employees' value-in-context may be self-affirmation gained from community engagement. The stakeholder value-in-context can contribute to the company's value-in-context in terms of reputation or positive word of mouth. Identifying the cognitive distance between the company and its stakeholders allows the company to enhance stakeholder value and improve the stakeholders' well-being through CSI.
- **Decision Metrics**: This paper offers tailored metrics to evaluate CSR practice from the perspective of stakeholders, including "know-what",

"know-how", and "know-who". "Know-what" is to examine the frequency of specific CSR dimensions mentioned by stakeholders, "know-how" is to examine the attitude and expectations of stakeholders towards CSR performance, and "know-who" is to identify the important stakeholders of certain CSR issues.

6.2 Conclusion

6.2.1 Summary of this Research

Big data has been viewed as a critical driver of industrial revolutions in recent years. Organisations embracing a data-transformation approach were found to shift towards a more service-centric standpoint. A data-driven service systems (DDSS) perspective is suggested to improve the understanding of how companies can facilitate big data and analytics to ameliorate the well-being of service systems. **Chapter 1** discusses the research background of DDSS and clarifies the research aim: to provide advanced knowledge of DDSS that allows researchers and practitioners to improve their capabilities of distilling "value" from big data.

This research is positioned as data analytics and business intelligence research, focusing on transforming big data into actionable insights to address business problems. A design science research approach is suitable for such data analytics research. **Chapter 2** justifies and details the use of design science to develop the DDSS framework. The DDSS framework, as the design artefact, consists of three management blocks: service systems, data environment, and decision-making. Within the three blocks, nine operational components are identified: actors, resources, activities, time, context, big data, big data analytics (BDA), value, and decision metrics.

To validate the DDSS framework, social media dialogue data is used, considering that social media are widely adopted by today's companies for various business purposes and that user interactions on social media form complex service ecosystems. Moreover, the data generated on social media is one of the most important sources of big data. A growing body of research has been conducted to examine this data. Three applications based on social media dialogue data are provided in the format of journal manuscripts in **Chapters 3**, **4**, and **5**.

Chapter 3 examines the social media dialogues during service recovery on Twitter customer care. This application aims to address the issue related to the dynamic factors affecting service recovery outcomes. The dynamic factors were assessed based on the DDSS operational components, especially actor (e.g. multiple-actor interaction, customer co-recovery) and time (e.g. interaction-based recovery process). The BDA techniques used in this application included sentiment analysis, text mining and logistic regression modelling. Finally, the value in this application was investigated by the decision metrics regarding better complaint management and customer post-recovery satisfaction.

Chapter 4 is a follow-up research of the first application, but this application is specially centred on BDA – developing a dialogue-mining framework to address the issues of customer recovery and process recovery. The dialogue-mining framework investigated three dimensions of dialogues – namely linguistic and semantic, process, and relationship dimensions – and it was tested using text mining and process mining. In addition to better complaint handling, this application offers an improved data analytical method for both the company and social media analytics.

Chapter 5 is conducted on a different research context – corporate social innovation (CSI), exploring the dialogues between a company and its stakeholders.

This application theorises and empirically tests a data-driven approach to supporting CSI. For the dialogue data, text mining was employed to extract stakeholder value regarding stakeholder "know-what", "know-who", "know-how", and "know-why" towards specific CSR activities. Then, a decision tree model was applied to examine how stakeholder value could predict CSI ideas. This application makes contributions to data-driven CSI and offers a set of propositions for practitioners to implement data-driven CSI.

Following the design research method, **Chapter 6** presents an in-depth discussion on the evaluation of the DDSS framework regarding the relevancy to the research environment, extensibility to the knowledge base, and utility for the field problems. An overall conclusion, research implications, and limitations are also provided in the final chapter to indicate directions for future research.

6.2.2 Contributions

This thesis explores "value" of big data and makes several contributions to both theoretical and managerial knowledge. Four research objectives specified in **Chapter 1** have been achieved: constructing a framework of DDSS (the first and second objectives), offering operational guidance including data acquisition, data exploitation and data-driven decision-making (the third objective), and demonstrating the utility of the DDSS framework in aiding problem-solving (the fourth objective).

In relation to the first and second research objectives, the DDSS framework clarifies the definition and builds linkages among service systems, big data, and decision-making. Advanced digital technologies and platforms accelerate the growth of big data and call for more specialised big data analytics (BDA) and data science research in the data market. However, big data research in marketing and service science is still a working-in-progress area. Prior research has approached issues such as data-enabling service revolutions (Rust and Huang 2014) and bridging the gap between data-driven marketing and traditional marketing strategies (Kumar et al. 2013; Wedel and Kannan 2016). However, when big data is viewed as an important driver to create, change, and improve service systems, a holistic framework that explains value creation mechanisms is still absent.

This thesis bridges the gap and conceptualises DDSS as special service systems relying on the use of big data to enhance value co-creation amongst actors. Drawing on a meso-system perspective, two types of value co-creation mechanisms are discussed: co-creation amongst actors in the service systems and co-creation between BDA and companies in the data environment. The former is achieved when companies' products, services, and processes (value propositions) are used by customers, or when customers' complaints (value propositions) are applied to improve companies' practices. The latter is achieved when companies facilitate BDA to transform big data into new value propositions such as service innovation. The DDSS framework highlights the importance of managing these two mechanisms of value co-creation to improve the well-being of system actors.

In relation to the third research objective, this thesis entails a set of operational components in the DDSS framework, which have an impact on tactical business operations. These components can serve as a toolkit allowing companies to manage the process of data acquisition, data exploitation, and data-driven decision-making. In the data acquisition stage, companies translate business problems into data analytical problems and investigate the components within service systems (e.g. actor, resource, activity, time, and context). This helps companies to define the representative data sources and unit of analysis. In the data exploitation stage, companies seek tailored BDA to deal with specific big data in the data environment. Several BDA techniques have already been in use in marketing, such as data and text mining, machine learning, and Bayesian methods (Rust and Huang 2014). It is worth noting that BDA techniques should be performed following relevant marketing metrics. Kumar et al. (2013) pointed out that data is only useful when it can inform metrics, such as customer acquisition, customer satisfaction, or share of wallet (SOW). The DDSS framework highlights the importance of clarifying the relationship between value and metrics for decision-making. In this way, it can help practitioners to apply data-driven insights to promote better value co-creation amongst focal actors within service systems.

The third contribution of the DDSS framework is its high applicability and usefulness in aiding problem-solving. The three papers provided in this thesis demonstrate how the proposed framework is implemented in terms of understanding the value co-creation mechanisms of field problems, capturing and analysing big data, and offering insights into the research problems. Table 6.5 summaries the key theoretical implications in each paper and clarifies how the findings from the three papers improve service literature. These papers present examples of novel applications of using text-mining analysers on Twitter data to investigate problems regarding service recovery and corporate social innovation. Each paper provides a clear framework that transforms business problems into data problems and extracts data-driven insights to improve business practices. The DDSS framework, as well as the frameworks in the papers, can serve as open learning models that are continuously advanced by follow-up research in the specific domains.

Table 6.5 Theoretical Implications of the Three papers

	Key Findings	Contributions to Prior Research
1st Paper	• Examined three dynamic components that influence service recovery performance. • Competitor involvement has a significant negative impact on the service-recovery experience. • Other users' engagement is a key moderator of service recovery activities and outcomes. • The increasing engagement of customers in service recovery is detrimental to service recovery. • A high level of interpersonal justice via Twitter (e.g. showing empathy and apology) would not promote positive service recovery experience. • Procedural justice (e.g. Follow-up) is the key to positive recovery experience.	• Demonstrated the use of Grönroos and Voima's (2013) value creation sphere in the service-recovery context. • Improved the understanding of how social influences (e,g, Schaefers and Schamari 2016) affect service recovery outcomes. • Improve the understanding of customer co-recovery (e.g. Dong et al. 2008) in the social media context. • Re-examined justice theory (e.g. Smith et al. 1999; Tax et al. 1998) using a new type of customer care platform.
2nd Paper	• Introduced a dialogue-mining framework, specifying three dimensions of dialogues: linguistic and semantic, process, and relationship. • The results of the process-recovery modelling highlighted the weak touchpoints during customer experience. • The second application showed the real path of customer recovery delivered by the company, thus, uncovering loops and inefficient processes. • The second application re-built the cross-channel associations during complaint handling, thus,	• Demonstrated the application of Grönroos (2004)'s marketing communication model in designing a data analytical pipeline • Demonstrated the application of customer journey (e.g. Lemon and Verhoef 2016) in aiding to improve process recovery • Extended Fan and Niu (2016)'s research scope by adding the temporal and contextual dimensions in analysing service-recovery dialogues.

	improving business resources allocation.	
3rd Paper	• Explored the use of data-driven approach in enhancing corporate social innovation. • Suggested a novel analytical method to unearth cognitive distance which hinders value co-creation	• Applied value co-creation literature (Lusch and Nambisan 2015; Prahalad and Ramaswamy 2004) in a rarely-examined context – corporate social innovation. • Clarified the mechanism of value co-creation between the company and the stakeholders and added to the co-created social innovation (Herrera 2015; Mirvis et al. 2016).
	• Proposed five propositions that deliver testable knowledge of corporate social innovation.	• Tested Herrera's (2015) stakeholder knowledge model (know-what, know-how, know-who and know-why) in driving corporate social innovation.

6.2.3 Limitations of this Research

While this thesis contributes to the improved understanding of DDSS in terms of theoretical foundations, practical applications, and data analytical methods, it is not without challenges. In the papers provided in **Chapters 3**, **4**, and **5**, the limitations of research have been disclosed, and directions for further research were suggested. In this section, the discussion on research limitations will focus on the design of the DDSS framework and the possible conflicts between the use of the DDSS framework and business practice.

The first limitation of the DDSS framework is related to its design. The DDSS framework can serve as a conceptual framework, a management toolkit, and a prototype of data analytics (e.g. ontologies), building linkages amongst theories, data analytics, and practices. However, it may not be optimised as the best possible combination of operational components, as the components were constructed at a high level of abstraction. Instead, the DDSS framework offers an improvement over existing models that were centred only on data sources and data analytics for addressing marketing problems (e.g. Kumar et al. 2013). As stressed in Gregor and Hevner (2013), evaluating improvement contributions is challenging and requires a thorough understanding of field problems, existing solutions, and available alternatives. Though the DDSS framework has been validated in three applications, a lack of practitioners' perspective (e.g. frontline employees, managers, BDA vendors) during the design process challenges the research outputs. This is because the DDSS framework is designed to address the wicked problems to which new digital technologies, new data sources and analytics are continuously being introduced in the data market and changing marketing activities. Therefore, the DDSS framework should be viewed as an "open learning" model (Ordenes et al. 2014), which can be enhanced and expanded to account for the changing environment over time.

Second, the DDSS framework was developed following service-dominant logic (SDL) to depict actor value co-creation within service systems. However, in the three applications, service logic (SL) and customer-dominant logic (CDL) were also employed to interpret the mechanisms of value creation of the field problems. More specifically, **Chapter 3** used SL to construct a dynamic service recovery framework, and **Chapter 4** explained the customer experience processes using CDL. Although the differences amongst these three research streams have been briefly discussed, the conflicts amongst SDL, SL, and CDL in applications of the DDSS framework are not thoroughly examined.

Third, the DDSS framework suggests the use of big data to improve value co-creation amongst actors in the service systems and enhance the well-being of the systems. However, some possible conflicts between the DDSS perspective and business practice should be identified as limitations. One of the most important limitations is the ethical concerns of using data. Big data is viewed as shared resources that companies provide through interaction platforms and customers generating data (Xie et al. 2016). Today's companies tend to develop a 360° view of customers through collecting more customer-centric data from various channels, such as customer profiles, social connections, and clickstreams (Kunz et al. 2017). Also, more and more new business models are enabled by the use of big data, such as using data to improve practices, digitising physical assets, trading data, and codifying a distinctive service capability (Parmar et al. 2014). However, customers may view such data collection and data use as a threat to their privacy and security (Kumar et al. 2013). With the potential unethical conduct of using consumer data, the DDSS perspective may harm customers' well-being rather than benefitting them.

6.2.4 Future Research

The limitations of the DDSS framework create new avenues for future research. More research is needed to explain the complex and dynamic interrelationships amongst the three management blocks of the DDSS framework. Notably, DDSS does not function based on a linear or cyclical process. The interruption in DDSS can start from any point in the three management blocks of service systems, data environment, and decision-making. Taking the advance of cloud computing as an example, the change started in data environment and led companies to offer software-as-a-service applications (e.g. Microsoft Office) as a replacement for traditional product-centric software. Finally, the new services influenced consumers' product purchase and use in the service systems (e.g. using applications of Microsoft Office on mobile devices).

Future research should engage the focal actors within service systems (e.g. customers, service agents), big data environment (e.g. data scientists, BDA vendors), and the decision-making process (e.g. managers) to obtain diverse worldviews of DDSS. This is especially important as the proliferation of new digital technology (e.g. artificial intelligence), new applications (e.g. virtual reality, augmented reality), and new data sources is expected to transform the business environment and lead to industry revolutions more quickly and drastically than before. The engagement of the aforementioned actors enables researchers to define the boundary of DDSS and examine various mechanisms of value co-creation.

In terms of the potential conflicts between the DDSS framework and business practice, further research should investigate the ethical concerns about using customer data to drive business practice and enhance existing business models. In particular, the use of customer data is not always related to enhancing customer value but rather to promoting value co-creation between a company and other system actors (e.g. selling customer data to third parties). Therefore, more research is suggested to analyse the research issues, including the ethical issues and the potential value co-destruction issues of using big data.

Overall, this thesis has delivered a novel artefact, with three empirical applications conducted on social media dialogue data. The applications demonstrate the usefulness in understanding the DDSS in the selected domains. Nevertheless, the potential of DDSS in marketing practice is still largely unexplored. The DDSS framework can serve as a reference point for further big data research in both academic and practical marketing. Several sub-disciplines of marketing, including e-commerce, social media customer care, content marketing, referral marketing, search engine marketing, and customer journey planning, are highly relevant to the DDSS perspective. It is suggested that future research can use and test the DDSS framework in the aforementioned areas to address relevant marketing problems and advance the domain knowledge.

References

van Aken, J.E. (2005), "Management Research as a Design Science: Articulating the Research Products of Mode 2 Knowledge Production in Management," *British Journal of Management*, 16(1), 19–36.

Ambler, T. (2003), *Marketing and the Bottom Line: The Marketing Metrics to Pump Up Cash Flow.* 2nd ed., London: Financial Times Prentice Hall.

Baird, C.H. and Parasnis, G. (2011), "From Social Media to Social Customer Relationship Management," *Strategy & Leadership*, 39(5), 30–37.

Beloglazov, A., Banerjee, D., Hartman, A. and Buyya, R. (2015), "Improving Productivity in Design and Development of Information Technology (IT) Service Delivery Simulation Models," *Jounral of Service Resesarch*, 18(1), 75–89.

Boinepelli, H. (2015), "Applications of Big Data," In H. Mohanty, B. Prachet, and D. Chenthati, eds., *Big Data: A Primer*. New Delhi: Springer.

Bradlow, E.T., Gangwar, M., Kopalle, P. and Voleti, S. (2017), "The Role of Big Data and Predictive Analytics in Retailing," *Journal of Retailing*, 93(1), 79–95.

Buchanan, R. (1992), "Wicked Problems in Design Thinking," *Design Issues. The MIT Press*, 8(2), 5–21.

Carillo, K.D.A. (2017), "Let's Stop Trying to be "Sexy" – Preparing Managers for the (Big) Data-driven Business Era," *Business Process Management Journal*, 23(3), 598–622.

Chandler, J.D. and Vargo, S.L. (2011), "Contextualization and Value-in-Context: How Context Frames Exchange," *Marketing Theory*, 11(1), 35–49.

Chen, C.L.P, and Zhang, C.Y. (2014), "Data-Intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data," *Information Sciences*, 275(10), 314–347.

Chen, H., Chiang, R.H.L. and Storey, V.C. (2012), "Business Intelligence and Analytics: From Big Data To Big Impact," *Mis Quarterly*, 36(4), 1165–1188.

Chen, Y., Kreulen, J., Campbell, M, and Abrams, C. (2011), "Analytics Ecosystem Transformation: A Force for Business Model Innovation," In *Proceedings of the 2011 Annual SRII Global Conference (SRII 2011)*, 11–20.

Chua, A.Y.K. and Banerjee, S. (2013), "Customer Knowledge Management via Social Media: The Case of Starbucks," *Journal of Knowledge Management*, 17(2), 237–249.

Côrte-Real, N., Oliveira, T. and Ruivo, P. (2016), "Assessing Business Value of Big Data Analytics in European Firms," *Journal of Business Research*, 70, 379–390.

Costello, L., Mcdermott, M.-L. and Wallace, R. (2017), "Netnography: Range of Practices, Misperceptions, and Missed Opportunities," *International Journal of Qualitative Methods*, 16, 1–12.

Davenport, T.H. and Patil, D.J. (2012), "Data Scientist: The Sexiest Job of the 21st Century," *Harvard Business Review*, 90(10), 70–76.

Echeverri, P. and Skalen, P. (2011), "Co-Creation and Co-Destruction: A Practice-Theory Based Study of Interactive Value Formation," *Marketing Theory*, 11(3), 351–373.

Erevelles, S., Fukawa, N. and Swayne, L. (2016), "Big Data Consumer Analytics and the Transformation of Marketing," *Journal of Business Research*, 69(2), 897–904.

European commission (2014), *Towards a Thriving Data-driven Economy*, COM(2014) 442. [Online] Avaliable at: https://ec.europa.eu/digital-single-market/news/communication-data-driven-economy.

Fan, W. and Gordon, M.D. (2014), "The Power of Social Media Analytics," *Communications of the ACM*, 57(6), 74–81.

Farris, P.W., Bendle, N.T., Pfeifer, P.E. and Reibstein, D.J. (2010), *Marketing Metrics:The Definitive Guide to Measuring Marketing Performance*, 2nd ed., New York, US: Pearson Education.

Feinleib, D. (2014), *Big Data Bootcamp: What Managers Need to Know to Profit from the Big Data Revolution*. Apress

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D. (2015), "How "Big Data" can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study," *International Journal of Production Economics*, 165, 234–246.

Gandomi, A. and Haider, M. (2015), "Beyond the Hype: Big Data Concepts, Methods, and Analytics," *International Journal of Information Management*, 35(2), 137–144.

Gantz, J. and Reinsel, D. (2011), “Extracting Value from Chaos State of the Universe: An Executive Summary,” *IDC iView*, (June), 1–12. [Online] Retrieved from: https://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf

George, G., Haas, M.R. and Pentland, A. (2014), “Big Data and Management,” *Academy of Management Journal*, 57(2), 321–326.

Goes, P.B. (2014), “Big Data and IS Research,” *MIS Quarterly*, 38(3), iii–viii.

Gregor, S. and Hevner, A.R. (2013), “Positioning and Presenting Design Science Research for Maximum Impact,” *MIS Quarterly*, 37(2), 337–355.

Grönroos, C. (2011), “Value Co-creation in Service Logic: A Critical Analysis,” *Marketing Theory*, 11(3), 279–301.

Grönroos, C. (2008), “Service Logic Revisited: Who Creates Value? And Who Co-creates?” *European Business Review*, 20(4), 298–314.

Grönroos, C. (1988), “Service Quality: The Six Criteria of Good Perceived Service Quality,” *Review of Business*, 9(3), 10–13.

Grönroos, C. and Gummerus, J. (2014), “The Service Revolution and its Marketing Implications: Service Logic v.s. Service-Dominant Logic,” *Managing Service Quality*, 24(3), 206–229.

Grönroos, C. and Voima, P. (2013), “Critical Service Logic: Making Sense of Value Creation and Co-Creation,” *Journal of the Academy of Marketing Science*, 41(2), 133–150.

Hanna, R., Rohm, A. and Crittenden, V.L. (2011), “We’re All Connected: The Power of the Social Media Ecosystem,” *Business Horizons*, 54(3), 265–273.

Hanssens, D.M. and Pauwels, K.H. (2016), “Demonstrating the Value of Marketing,” *Journal of Marketing*, 80(6), 173–190.

Harris, D. (2012), *Netflix Analyzes a lot of Data about Your Viewing Habits. GigaOM*. Available at: https://gigaom.com/2012/06/14/netflix-analyzes-a-lot-of-data-about-your-viewing-habits/.

Hartmann, P.M., Zaki, M., Feldmann, N. and Neely, A. (2016), “Capturing Value from Big Data – A Taxonomy of Data-Driven Business Models Used by Start-up Firms,” *International Journal of Operations & Production Management*, 36(10), 1382–1406.

He, W., Wu, H., Yan, G., Akula, V. and Shen, J. (2015), "A Novel Social Media Competitive Analytics Framework with Sentiment Benchmarks," *Information and Management*, 52(7), 801–812.

Heinonen, K., Strandvik, T., Mickelsson, K-J., Edvardsson, B., Sundström, E., and Andersson, P. (2010), "A Customer-Dominant Logic of Service," *Journal of Service Management*, 21(4), 531–548.

Heinonen, K. and Strandvik, T. (2015), "Customer-Dominant Logic: Foundations and Implications," *Journal of Services Marketing*, 29(6/7), 472–484.

Heinonen, K., Strandvik, T. and Voima, P. (2013), "Customer Dominant Value Formation in Service," *European Business Review*, 25(2), 104–123.

Hevner, A.R. (2007), "A Three Cycle View of Design Science Research," *Scandinavian Journal of Information Systems*, 19(2), 87–92.

Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004), "Design Science in Information Systems Research," *MIS Quarterly*, 28(1), 75–105.

Hiroshi, I. (2015), *Social Big Data Mining*, CRC Press.

Hofer-Shall, Z. (2010), *The Forrester Wave: Listening Platforms, Q3 2010*, Forrester Research

Hoffman, D.D.L.D.L. and Fodor, M. (2010), "Can You Measure the ROI of Your Social Media Marketing?" *MIT Sloan Management Review*, 52(1), 41–49.

Huang, M.-H. and Rust, R.T. (2013), "IT-Related Service: A Multidisciplinary Perspective," *Journal of Service Research*, 16(3), 251–258.

IBM, (2016), *Getting to Know You: How Leading Marketers Are Using Social Media Analytics*, [Online] Available at: https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=ASW12375USEN.

Kaplan, A.M. and Haenlein, M. (2010), "Users of the World, Unite! The Challenges and Opportunities of Social Media," *Business Horizons*, 53(1), 59–68.

Katona, Z., Zubcsek, P.P. and Sarvary, M. (2011), "Network Effects and Personal Influences: The Diffusion of an Online Social Network." *Journal of Marketing Research*, 48(3), 425–443.

Khan, G.F. (2015), *Seven Layers of Social Media Analytics: Mining Business Insights from Social Media Text, Actions, Networks, Hyperlinks, Apps, Search Engine, and Location Data*, CreateSpace Independent Publishing Platform.

Kiron, D., Prentice, P.K. and Ferguson, R.B. (2012), "Innovating with Analytics," *MIT Sloan Management Review*, 54(1), 47-52.

Kozinets, R. V. (2010), *Netnography: Doing Ethnographic Research Online*, SAGE Publications.

Kozinets, R. V. (2002), “The Field Behind the Screen: Using Netnography for Marketing Research in Online Communities,” *Journal of Marketing Research*, 39(1), 61–72.

Kumar, V., Chattaraman, V., Neghina, C., Skiera, B., Aksoy, L., Buoye, A., and Henseler, J. (2013), “Data-Driven Services Marketing in A Connected World,” *Journal of Service Management*, 24(3), 330–352.

Kumar, V. (2016), “Introduction: Is Customer Satisfaction (Ir)relevant as a Metric?” *Journal of Marketing*, 80(5), 108–109.

Kumar, V. and Reinartz, W. (2006), *Customer Relationship Management: A Databased Approach*, NY: John Wiley & Sons.

Kunz, W., Aksoy, L., Bart, Y., Heinonen, K., Kabadayi, S., Ordenes, F. V., Sigala, M., Diaz, D. and Theodoulidis, B. (2017), “Customer Engagement in a Big Data World,” *Journal of Services Marketing*, 31(2), 161–171.

Kuzgun, E. and Asugman, G. (2015), “Value in Services – A Service Dominant Logic Perspective,” *Procedia - Social and Behavioral Sciences*, 207, 242–251.

Lavalle, S., Lesser, E., Shockley, R., Hopkins, M. S. and Kruschwitz, N. (2011), “Big Data, Analytics and the Path From Insights to Value,” *MIT Sloan Management Review*, 52(2), 21–32.

Leroi-Werelds, S., Streukens, S., Brady, M. K. and Swinnen, G. (2014), “Assessing the Value of Commonly Used Methods for Measuring Customer Value: A Multi-Setting Empirical Study,” *Journal of the Academy of Marketing Science*, 42(4), 430–451.

Löbler, H., (2013), “Service-Dominant Networks: An Evolution from the Service-Dominant Logic Perspective,” *Journal of Service Management*, 24(4), 420–434.

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

Lusch, R.F., Vargo, S.L. and Tanniru, M. (2010), “Service Value Networks and Learning,” *Journal of the Academy of Marketing Science*, 38(1),19–31.

Maglio, P.P. and Spohrer, J. (2008), “Fundamentals of Service Science,” *Journal of the Academy of Marketing Science*, 36(1), 18–20.

Malthouse, E.C. and Li, H. (2017), "Opportunities for and Pitfalls of Using Big Data in Advertising Research," *Journal of Advertising*, 46(2), 227–235.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A.H. (2011), *Big Data: The Next Frontier for Innovation, Competition and Productivity*, McKinsey Global Institute. [Online] Avaliable at: https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation.

March, S.T. and Smith, G.F. (1995), "Design and Natural Science Research on Information Technology," *Decision Support Systems*, 15(4), 251–266.

McAfee, A. and Brynjolfsson, E. (2012), "Big Data. The Management Revolution," *Harvard Buiness Review*, 90(10), 61–68.

Michel, S., Bowen, D. and Johnston, R. (2009), "Why Service Recovery Fails: Tensions among Customer, Employee, and Process Perspectives," *Journal of Service Management*, 20(3), 253–273.

Mickelsson, K-J. (2013), "Customer Activity in Service," *Journal of Service Management*, 24(5), 534–552.

Mintz, O. and Currim, I.S. (2013), "What Drives Managerial Use of Marketing and Financial Metrics and does Metric Use Affect Performance of Marketing-Mix Activities?" *Journal of Marketing*, 77(March), 17–40.

Moe, W.W. and Schweidel, D.A. (2014), *Social Media Intelligence*, Cambridge: Cambridge University Press.

Moro, S., Rita, P. and Vala, B. (2016), "Predicting Social Media Performance Metrics and Evaluation of the Impact On Brand Building : A Data Mining Approach," *Journal of Business Research*, 69(9), 3341–3351.

Nadkarni, A. and Vesset, D. (2015), *Worldwide Big Data Technology and Services 2015 – 2019 Forecast*, Market Forecast, IDC Research.

Netzer, O., Feldman, R., Goldenberg, J., and Fresko, M. (2012), "Mine Your Own Business: Market-Structure Surveillance Through Text Mining," *Marketing Science*, 31(3), 521–543.

van Noort, G., Willemsen, L. M., Kerkhof, P. and Verhoeven, J.W.M. (2014), "Webcare as An Integrative Tool for Customer Care, Reputation Management and Online Marketing: A Literature Review," In P. J. Kitchen and E. Uzunoglu, eds., *Integrated Communications in the Postmodern Era*, Springer, 77–99.

Normann, R. and Ramírez, R. (1993), “From Value Chain to Value Constellation: Designing Interactive Strategy,” *Harvard Business Review*, 71(4), 65–77.

Ordenes, F. V., Theodoulidis, B., Burton, J., Gruber, T. and Zaki, M. (2014), “Analyzing Customer Experience Feedback Using Text Mining: A Linguistics-Based Approach,” *Journal of Service Research*, 17(3), 278–295.

Ordenes, F. V., Ludwig, S., de Ruyter, K., Grewal, D. and Wetzels, M. (2017), “Unveiling What is Written in the Stars: Analyzing Explicit, Implicit and Discourse Patterns of Sentiment in Social Media,” *Journal of Consumer Research*, 43(6), 875–894.

Osterwalder, A. (2004), “ The Business Model Ontology A Proposition in A Design Science Approach,” Ph.D thesis, Université de Lausanne, Switzerland.

Parmar, R., Mackenzie, I., Cohn, D. and Gann, D. (2014), “The New Patterns of Innovation,” *Harvard Business Review,* 86-97

Patricio, L., Fisk, R. P., Falcao e Cunha, J., and Constantine, L. (2011), “Multilevel Service Design: From Customer Value Constellation to Service Experience Blueprinting,” *Journal of Service Research*, 14(2), 180–200.

Plé, L. and Cáceres, R.C. (2010), “Not Always Co-Creation: Introducing Interactional Co-Destruction of Value in Service-Dominant Logic,” *Journal of Services Marketing*, 24(6), 430–437.

Prahalad, C.K. and Ramaswamy, V. (2000), “Co-Opting Customer Competence,” *Harvard Business Review*, 78(1), 79–90.

Prahalad, C.K. and Ramaswamy, V. (2004), “Co-Creation Experiences: The Next Practice in Value Creation,” *Journal of Interactive Marketing*, 18(3), 5–14.

Ratliff, J.M. and Kunz, M.B. (2014), “A Profile of Social Media Use By Fortune 500 and Inc. 500 Corporations,” *Academy of Strategic Management Journal*, 13(2), 61–73.

Ruhi, U. (2014), “Social Media Analytics as A Business Intelligence Practice: Current Landscape & Future Prospects,” *Journal of Internet Social Networking & Virtual Communities*, (2012), 1–13.

Rust, R.T. and Huang, M.-H. (2014), “The Service Revolution and the Transformation of Marketing Science,” *Marketing Science*, 33(2), 206–221.

Samantha, N. and Pilecki, M. (2016), *The Forrester Wave: Enterprise Social Listening Platforms, Q1 2016,* Forrester Research

Schaefers, T. and Schamari, J. (2016), "Service Recovery via Social Media: The Social Influence Effects of Virtual Presence," *Journal of Service Research*, 19(2), 192–208.

Simon, H.A. (1996), *The sciences of the artificial*, 3rd ed., Cambridge, UK: MIT Press.

Singaraju, S.P., Nguyen, Q. A., Niininen, O., and Sullivan-Mort, G. (2016), "Social Media and Value Co-creation in Multi-Stakeholder Systems: A Resource Integration Approach," *Industrial Marketing Management*, 54, 44–55.

Stieglitz, S. and Dang-Xuan, L. (2013), "Social Media and Political Communication: A Social Media Analytics Framework," *Social Network Analysis and Mining*, 3(4), 1277–1291.

Tambe, P. (2014), "Big Data Investment, Skills, and Firm Value," *Management Science*, 60(6), 1452–1469.

Teixeira, J.G., Patrício, L., Huang, K-H., Fisk, R. P., Nóbrega, L. and Constantine, L. (2017), "The MINDS Method: Integrating Management and Interaction Design Perspectives for Service Design," *Journal of Service Research*, 20(3), 240–258

Tsang, E. and Kwan, K. (1999), "Replication and Theory Development in Organizational Science: A Critical Realist Perspective," *The Academy of Management Review*, 24(4), 759–780.

Tynan, C., McKechnie, S. and Hartley, S. (2014), "Interpreting Value in the Customer Service Experience Using Customer-Dominant Logic," *Journal of Marketing Management*, 30(9-10), 1058–1081.

Vaishnavi, V.K. and Kuechler, W. (2015), *Design Science Research Methods and Patterns*, 2nd ed., CRC Press.

Vargo, S.L. and Akaka, M.A. (2012), "Value Cocreation and Service Systems (Re)Formation: A Service Ecosystems View," *Service Science*, 4(3), 207–217.

Vargo, S.L. and Lusch, R.F. (2004), "Evolving to a New Dominant Logic for Marketing," *Journal of Marketing*, 68(1), 1–17.

Vargo, S.L. and Lusch, R.F. (2008), "Service-Dominant Logic: Continuing the Evolution," *Journal of the Academy of Marketing Science*, 36(1), 1–10.

Vargo, S.L., Maglio, P.P. and Akaka, M.A. (2008), "On Value and Value Co-creation: A Service Systems and Service Logic Perspective," *European Management Journal*, 26(3),145–152.

Vargo, S.L. and Lusch, R.F. (2017), "Service-Dominant Logic 2025," *International Journal of Research in Marketing*, 34(1), 46–67.

van de Ven, A.H. (2007), *Engaged Scholarship: A Guide for Organizational and Social Research*. Oxford University Press.

Weber, E.P. and Khademian, A.M. (2008), "Wicked Problems, Knowledge Challenges, and Collaborative Capacity Builders in Network Settings," *Public Administration Review*, 68(2), 334–349.

Webster, F.E. and Lusch, R.F. (2013), "Elevating Marketing: Marketing is Dead! Long Live Marketing!" *Journal of the Academy of Marketing Science*, 41(4), 389–399.

Wedel, M. and Kannan, P.K. (2016), "Marketing Analytics for Data-Rich Environments," *Journal of Marketing*, 80(6), 97–121.

Weill, P. and Woerner, S.L. (2015), "Thriving in an Increasingly Digital Ecosystem," *MIT Sloan Management Review*, 56(4), 27–34.

Wilden, R., Akaka, M.A., Karpen, I.O. and Hohberger, J. (2017), "The Evolution and Prospects of Service- Dominant Logic: An Investigation of Past, Present, and Future Research," *Journal of Service Research*, In Press.

Wu, C., Buyya, R. and Ramamohanarao, K. (2016), "Big Data Analytics = Machine Learning + Cloud Computing," In R. Buyya, R. Calheiros, and A. Dastjerdi, eds., *Big Data: Principles and Paradigms*. Morgan Kaufmann, 3–38.

Xie, K., Wu, Y., Xiao, J. and Hu, Q. (2016), "Value Co-Creation between Firms and Customers: The Role of Big Data-Based Cooperative Assets," *Information & Management*. 53(8), 1034–1048.

Xu, Y., Marshall, R., Edvardsson, B., and Tronvoll, B. (2014), "Show You Care: Initiating Co-creation in Service Recovery," *Journal of Service Management*, 25(3), 369–387.

Zeng, D., Chen, H., Lusch, R. and Li, S-H. (2010), "Social Media Analytics and Intelligence," *IEEE Intelligent Systems*, 25(6), 13–16.

Zolnowski, A., Christiansen, T. and Gudat, J. (2016), "Business Model Transformation Patterns of Data-Driven Innovations," In the *Proceeding of 24th European Conference on Information Systems, ECIS 2016*, (June), 0–16.