공학박사학위논문

Development and Roadmapping of Smart Service Systems Using Data-driven and User Innovation Approach

데이터 기반 및 사용자 혁신 접근을 활용한

스마트 서비스 시스템의 개발 및 로드맵핑

2019 년 8 월

서울대학교 대학원

산업공학과

한 민 탁

Abstract

Development and Roadmapping of Smart Service Systems Using Data-driven and User Innovation Approach

Mintak Han

Department of Industrial Engineering

The Graduate School

Seoul National University

In many industries, the smart service system is of great significance in creating value in the interaction of users and products, technologies, and services through data. Moreover, the use of new or non-traditional data sources and method makes competitive advantages for firms. Quite naturally, firms face with this important question: how should we design a novel smart service systems and which values to be delivered to users? The answer, as most innovation research and practices do, is systematic service planning.

To this end, this thesis argues for new directions in the field of new smart service system innovation. Specifically, this study defines the distinctive characteristics of smart service systems; then deals with three research questions of service planning: system, technology, and data innovation through three research themes. For each research theme, several methodologies are revised and utilized to address each research question, providing the methodological sufficiency. A concrete

framework to effectively employ each methodology is descripted with detailed procedures.

The first study focuses on the research theme system innovation of a smart service system, and deals with the determination of components to develop new smart service concepts using morphological analysis. This study proposes a data-driven approach by incorporating mobile app service documents to increase objectivity and diversity in the construction of a morphology matrix. To this end, firstly, the novelty-quality map is developed to identify innovative data based on quantitative indicators. Secondly, morphological analysis is employed along with experts' judgment in order to generate new smart service concepts. Finally, the feasibility and effectiveness of the proposed approach are shown based on a comparative analysis with conventional approaches and real services through a case study of smart home.

The second study focuses on the research theme for technology innovation of a smart service system, and deals with making an improvement to the existing system using case-based reasoning (CBR). This study proposes a novel textual CBR approach which modified and integrated approach of SAO (Subject-Action-Object) analysis and technology tree into case representation and case retrieval process respectively. Specifically, by classifying problem and solution casebases focusing on SAO structures and searching for exact problem-solution sets related to the specific function of a product using technology tree, the approach enables the identification of various issues for incremental and even disruptive innovation. The effectiveness of the proposed approach is demonstrated through a case of drone technology using lead user communities.

Finally, the third study focuses on the research theme for data innovation of a smart service system. This study suggests a concept of data-integrated

technology roadmap, and proposes relevant structure, typology, and roadmapping process to show how to integrate data into the technology roadmapping. First, types of data integration are defined based on the literature review and practical business cases. Second, concept and structure of data-integrated technology roadmap is suggested, adding a data layer as intermediate and functional link for planning corresponding smart services. The data layer also consists of two sub-layers: internal data and external data. Third, typology of data-integrated technology roadmap is also suggested considering types of data integration. Based on the typology, this study also illustrates roadmapping processes for practical case of smart feeding machine.

In whole, this thesis can help to provide concrete and systematic tools that facilitate smart service system development by considering user innovation and big data approach. Ultimately it is expected to yield a foothold for the service designers and developers to proactively investigate innovation opportunities and candidates of feasible technologies by tapping into the massive and ever-growing collective knowledge from online user communities. Moreover, a suggested data-integrated technology roadmap enables service managers to visualize the development plans for a certain element of smart service systems to communicate effectively with colleagues.

Keywords: Smart service system development, Service innovation planning, User innovation, Data-driven approach, Data-integrated technology roadmap
Student Number: 2014-21804

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

Today, we are facing smart service systems such as smart home, smart car and smart health in every corner of our lives. Completely new smart service systems are being created every moment and intelligence level of the system has gradually increased [71, 123]. Given the drastic growth of the smart service systems, much attention has been paid to them in both academia and practice. Moreover, smart service system innovation is increasingly considered as a core activity for firms to maintain a competitive edge in the business environment [42].

For this reason, a substantial amount of research has been conducted regarding the definition [19, 90, 94, 109], characteristics [25, 82], and design [90] of "smart service systems". Among many research, one of the important findings is that the key elements of smart service systems are data and user [83, 90]. It differs from the existing concept of product-service system, which has primarily been studied in perspective of service innovation [5, 28, 32] and thus requires a systematic approach reflecting these distinctive characteristics for the development of smart service systems.

In more detail, rapid development of sensor technologies is making explosively variety forms of data which never existed before. At the same time, data technologies including storage, analytics, connectivity, and synchronization are evolving faster than ever. In this situation, there have been widespread views in the smart service system development process that data itself should be considered as a subject of management [25, 81, 83, 90]. It means that at the design stage, it is necessary to determine what kinds of data should be collected, processed, and delivered (i.e. input – process – output). The exchange and flow of these data take place in the interaction among products, services, and users. Taken together, the value that the smart service system delivers to the user is ultimately determined by the data innovation and the user's perception. Generally, the perception of users depends on the novelty and quality of functions that users expect [35]. Thus, proactively exploring which functions users want and devising a solution to address the issues enables incremental innovation to improve existing smart service systems and even disruptive innovation to discover entirely new smart service systems.

Yet despite the increasing importance of data and users in smart service systems, research on service innovation planning reflecting these distinctive and inherent nature is lacking. It leads to significant managerial issues in three ways. First, it becomes more difficult to design smart service systems because of increased component diversity. Second, unexpected issues can occur while using the smart service systems since the inter-relationships among the components make the system complex. Third, the generation and flow of data, which is an invisible object, increase the difficulty of managing the smart service system development. This thesis focuses on those problems arising due to the emerging concept of the smart service system in perspective of service innovation planning.

1.2 Research Objectives

The research objective of this thesis is to organize and apply the data-driven and user innovation approach for the development of smart service systems. From the theoretical perspective, this thesis tries to provide a solid framework and method to plan the service innovation, encompassing the broad range of planning processes. To this end, the smart service system is defined through extensive literature review and analysis of practical cases. From the methodological perspective, this thesis tries to fill the void in the literature related to the application of methodologies in service innovation planning. Specifically, this thesis is aimed at answering the following main questions.

(a) Which components to be considered in the structural aspect when designing a smart service system?

(b) What kinds of data sources are useful and which techniques are suitable for improving existing smart service systems or exploring new smart service systems?

(c) How could we plan and communicate smart service system concepts in an effective and efficient way?

Each question is addressed through three research themes, each of which related to each of three service innovation planning problems. The objectives of the three themes are as follows.

The first theme suggests a data-driven morphological analysis to determine the components of new smart service concepts. This study proposes a novelty-quality map by incorporating mobile app service documents in the construction of a morphology matrix. Using this map, we categorize service documents as one of four types – high innovative potential, universal service, out of competition, question mark – by relative novelty and quality score.

The second theme aims to explore potential problems and related feasible solutions by utilizing online lead user communities. This study proposes a novel textual CBR approach that modified and integrated approach of SAO (Subject-Action-Object) analysis and technology tree into case representation and case retrieval process respectively. Specifically, by classifying problem and solution casebases focusing on SAO structures and searching for exact problem-solution sets related to the specific function of a product using a technology tree, the approach enables the identification of various issues for incremental and even disruptive innovation.

Finally, the third theme suggests a concept of data-integrated technology roadmap, and proposes relevant structure, typology, and roadmapping process to show how to integrate data into the technology roadmapping. Based on the typology, this study also illustrates roadmapping processes for practical cases. A suggested data-integrated technology roadmap enables service managers to visualize the development plans for a certain element of smart service systems to communicate effectively with colleagues.

1.3 Scope and Framework

This thesis defines two distinctive characteristics of smart service systems as data innovation and user-driven development. It means that the data itself is subject to service innovation planning. Moreover, new values are created by interaction with the users. Thus, when developing any service, in addition to utilizing innovative technologies, a necessary condition is satisfying user needs. Consequently, it is required to explore potential problems based on user experience in the ideation stage for improvement of the smart service system.

Each aspect can cause significant managerial problems, and three of them are addressed in this research. Then three management issues related to data innovation and user-driven development are transformed into research themes as shown in Figure 1.1.

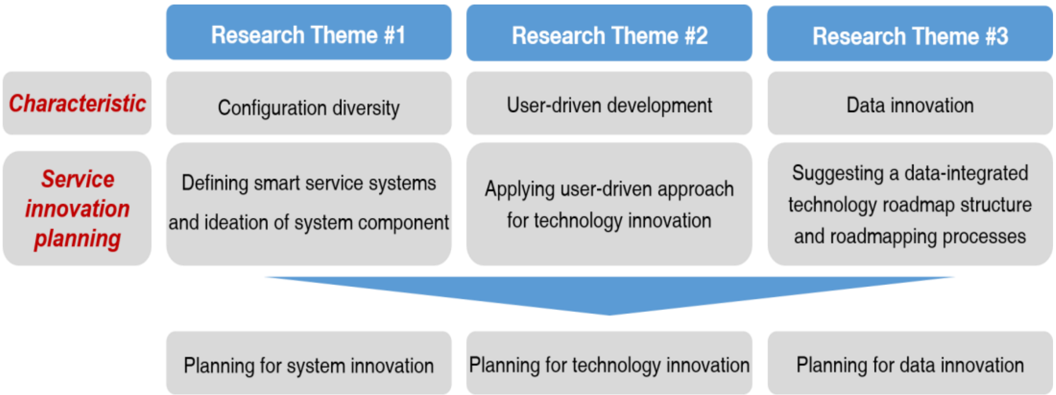

Figure 1.1 Service innovation planning issues of this thesis

The research themes derived from the previous process are as follows: (1) defining and designing smart service systems; (2) applying data-driven and user innovation approach for technology innovation; (3) suggesting a data-integrated technology roadmap structure and roadmapping process. Each theme is effectively addressed in this research by employing and creatively combining appropriate methods as shown in Figure 1.2.

Planning for system innovation	**Planning for technology innovation**	**Planning for data innovation**
Novelty detection technique	SAO semantic analysis	Technology roadmap
• The mechanism is able to identify incoming sensory pattern as being hitherto unknown	• SAO structure is a key concept that can show the relationship between components used in a relevant document	• A flexible planning technique to support strategic and long-range planning • Lack of attempts to consider the role of data in technology roadmap • Overall roadmapping process should be changed due to distinctive characteristics of smart service systems
Sentiment analysis	Technology tree	
• It is widely applied to voice of the customer materials such as reviews and survey responses, online and social media	• A hierarchical visual representation of the possible sequences of technologies	
Morphological analysis	Textual case-based reasoning	
• It is a method for identifying and investigating the total set of possible relationships contained in any given multi-dimensional problem	• It addresses the manipulation of cases expressed in text form	

Figure 1.2 The methods adopted in each chapter of this thesis

1.4 Organization of the Thesis

This thesis is composed of six chapters as illustrated in Figure 1.3. The remainder of this thesis is organized as follows. Chapter 2 provides background both from a theoretical and methodological perspective. The theoretical background was covered by reviewing the distinctive characteristics of smart service systems and online user communities in terms of user innovation. In this section, the definition and characteristics of smart service systems are discussed and the role of data and users in service innovation is also investigated. The methodological background covers a variety of methodologies and tools to support service innovation planning. Three leading methods (i.e. morphological analysis, textual case-based reasoning, and technology roadmap), which are adopted in this thesis are explained in detail, with their theoretical background and methodological strength.

Chapter 3, 4, 5 are main bodies of this thesis. As explained in the framework of this thesis, three research themes derived from the three innovation planning (i.e. system innovation, technology innovation, data innovation) are covered in these chapters respectively. According to the purpose of the theme, each theme encompasses its own introduction, background, appropriate methods with a practical case study and conclusions. Finally, this thesis ends with concluding remarks in Chapter 6.

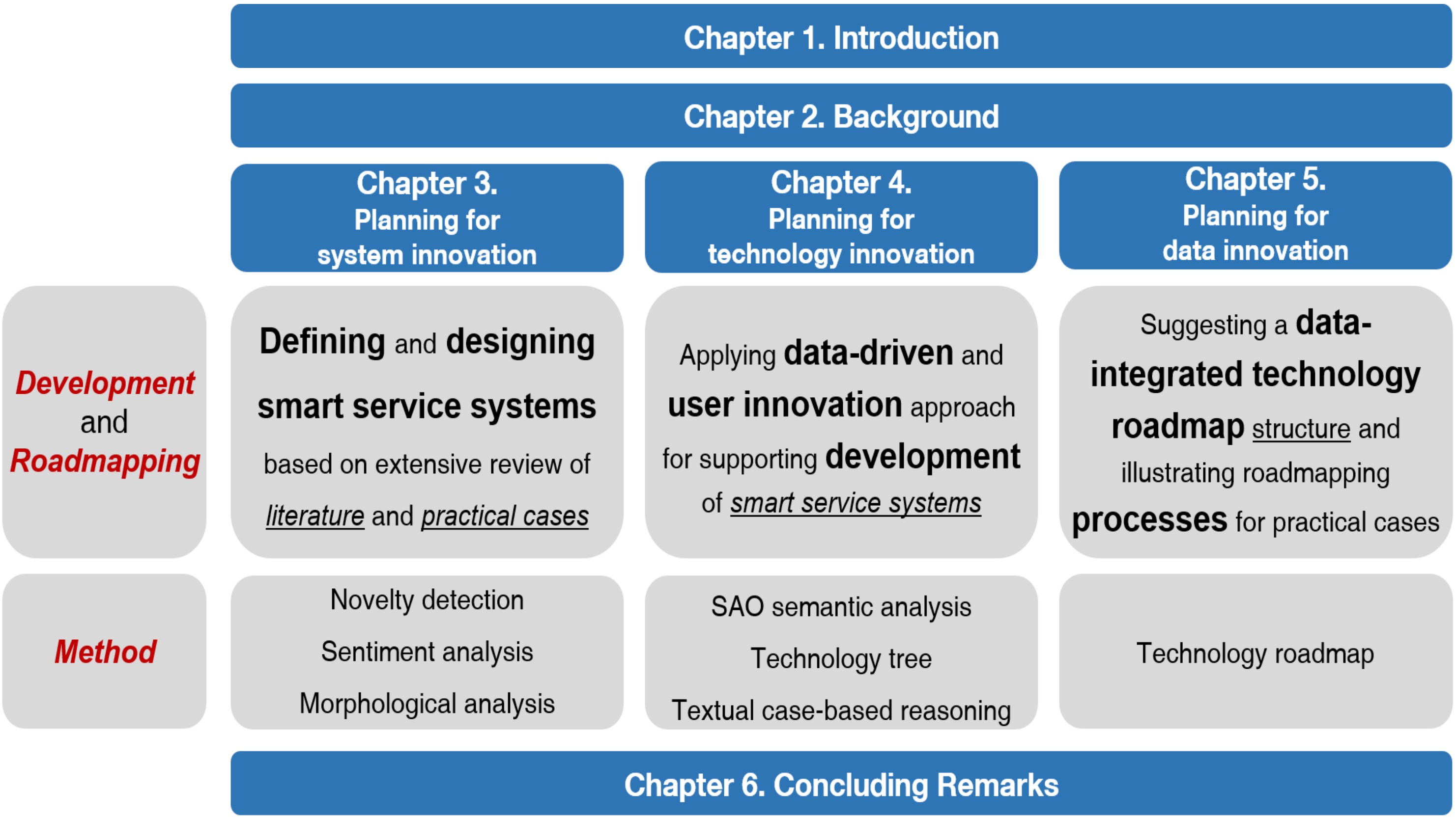

Figure 1.3 Overall structure of this thesis

Chapter 2

Background

2.1 Theoretical Background

2.1.1 Distinctive Characteristics of Smart Service Systems

For smart service system development, the prerequisite is to understand the smart service system itself. This section clarifies the conceptual backgrounds of smart service systems based on extensive literature reviews.

As service systems become increasingly smarter than ever with rapid technological development in many industries, researchers have investigated the smart service systems. Through extensive literature review, Table 2.1 lists several existing definitions, descriptions, and characteristics of smart service systems or smart product-service systems (PSSs). These definitions, descriptions, and characteristics are consistent in that they identify the capabilities and requirements of smart service systems. According to Lim and Maglio [82], it is reasonable to define a smart service system as a service system capable of learning, dynamic adaptation, and decision making [94] that requires an intelligent object [4, 127] and involves intensive data and information interactions among people and organizations [83, 90].

Table 2.1 Previous studies of smart service system concept

References	Definition/ Description/ Characteristic
Medina-Borja, 2015 [94]	A smart service system is "a service system capable of learning, dynamic adaptation, and decision making based upon data received, transmitted, and/or processed to improve its response to a future situation".
Maglio et al., 2015 [91]	Smart service systems are a kind of human-centered service system, meaning that knowledge, capabilities, and value are all determined by the people in the system.
Valencia et al., 2015 [113]	Smart PSS is the integration of smart products and e-services into single solutions delivered to the market to satisfy the needs of individual consumers.
Lerch and Gotsch, 2015 [79]	Manufacturers of digitalized PSS not only provide complex PSS to their customers, but also incorporate ICT solutions as a novel component in the product-service bundle, creating intelligent, independent operating systems that deliver the highest level of availability possible and optimize operations while reducing resource inputs.
Ostrom et al., 2015 [100]	In the new technology-enabled service context, customers increasingly create their own experiences in a more dynamic and autonomous way.
Spohrer and Demirkan, 2015 [109]	Smart service systems are ones that continuously improve (e.g., productivity, quality, compliance, sustainability, etc.) and co-evolve with all sectors. Because of analytics and cognitive systems, smart service systems adapt to a constantly changing environment to benefit customers and providers. Using big data analytics, service providers try to compete for customers by (1) improving existing

	offerings to customers, (2) innovating new types of offerings, (3) evolving their portfolio of offerings and making better recommendations to customers, (4) changing their relationships to suppliers and others in the ecosystem in ways their customers perceive as more sustainable, fair, or responsible.
Lim et al., 2016 [81]	Smart service systems are those service systems in which connected things and automation enable intensive data and information interactions among people and organizations that improve their decision making and operations. As the definition indicates, a smart service system consists of four components: (1) connected things, (2) automation, (3) people and organizations, and (4) data and information interactions.
Chowdhury et al., 2018 [25]	Smart PSS is the combinations and interactions between smart technologies, physical products, services, and business models. The interactions are essential to fulfill customers' needs.
Beverungen et al., 2019 [14]	Smart service is the application of specialized competences, through deeds, processes, and performances that are enabled by smart products. Smart service systems are service systems in which smart products are boundary-objects that integrate resources and activities of the involved actors for mutual benefit.

From a practical perspective, the integration of products and services is deeply associated with smart environment. with the development of Information and Communication Technologies (ICT), a variety of products are able to interact with each other and cooperate with other products to create new applications or services through wireless and wired connections. Such infrastructure has been named as 'smart environment'. Although there is no agreed definition of smart environment, it is reasonable to define smart environment as a global infrastructure for the information society, enabling advanced services by interconnecting things based on existing and evolving interoperable ICT in general [116]. Based on this, 'smart products' can be defined as things have identities and virtual personalities operating in smart environments using intelligent interfaces to connect and communicate within social, environment, and user contexts [111]. An integration of these smart products with services makes new values and diversified types of PSSs, ranging from a simple form such as an automated maintenance service added to products to a complex form such as smart home or smart grid [44, 78].

In the PSS within smart environment, because of integration with smart products, services that were not feasible in previous years have become possible and the range of service is gradually enlarged. That is, smart products can be recognized as a trigger for the development of new services [32]. In this process, 'data' serves as a core role. The smart product generates the new types of data, involving rich and spatial distributed identification, as well as historical and sensor data [33, 47]. Thus, a new service can be created and extended based on these huge amounts of data available in smart products [96]. For instance, U-health wearable devices provide health care service by delivering biometric information such as electrocardiogram and blood pressure to hospital.

Considering the conceptual background and phenomena, smart service systems have the following distinctive characteristics compared to the existing concept of PSS, as shown in Figure 2.1: (1) system components have been diversified, thus leading to increase the complexity of the system, (2) user interaction through smart products with ICT is inevitable for value creation, and (3) big data analytics has a crucial role in smart service system.

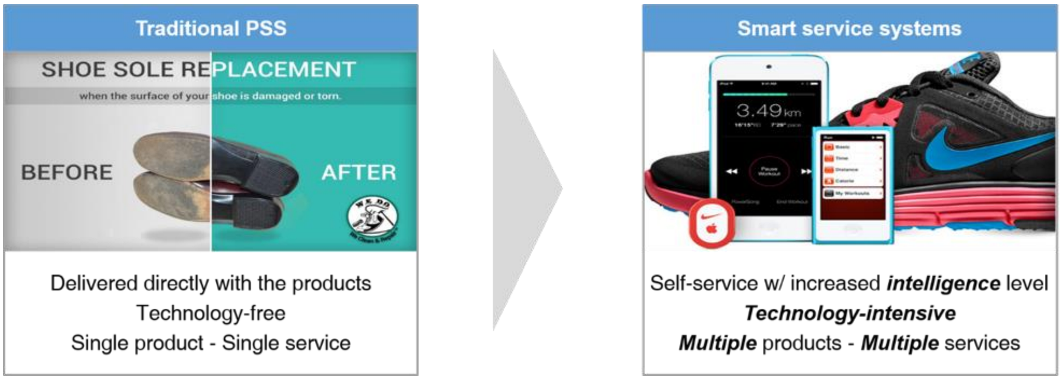

Figure 2.1 Stark differences from the existing PSS

Consequently, these distinctive characteristics enhance the importance of design and value creation, especially in the process of smart service system development. Table 2.2 shows the related features of smart service systems to be considered seriously in each process.

Table 2.2: Distinctive features of smart service systems

Aspect		Features of smart service systems
Design	Configuration	Smart products – delivery of smart service systems spanning the boundary among provider and customer
	Capability	Technologies – the intelligence level is determined by the complex work of various technologies such as sensors, embedded software, and analytical tools
Value creation	User interaction	Value is generated through understanding the context where a user utilize a smart service system and providing a proper functionality
	Data	Digital data-driven value creation and value capture for improving user relationship

2.1.2 User Innovation and Online User Communities

User innovation by individuals has moved from being considered peculiarities to being recognized as an important activity conducted by many users, which results in development of better products or services [15, 40, 119, 120]. Free and open source software is probably the most well-known case, where geographically dispersed individuals are able to develop software applications superior to other commercialized products [30]. As an example of consumer products, a variety of opinions from 197 end-users were gathered to develop extreme sports equipment [38].

As user innovation has become a strategic alternative in innovation management, user innovation approaches have also evolved in various ways. Particularly, one of the most interesting ways to user innovation is represented by the use of big data to provide novel insights for organizations [36]. With advances in information and communication technologies such as web 2.0, companies can gain access to new information, expertise, and ideas on online communities such as social media, brand community, and user innovation community [26, 61]. Online communities are a virtual organizational form in which knowledge collaboration can occur in unparalleled scale and scope, in ways not heretofore theorized [34]. The users in these communities tend to openly share their information, ideas, and even technologies with other members, exchanging product or service feedback and developing innovative ideas [65]. Indeed, it has been found through extensive research that a vast amount of these textual data can be a great source of innovation.

Various forms of online communities emerge due to advances in information and communication technologies such as web 2.0 [65] and thus heterogeneous contents are covered according to the characteristics of each community. In the following community types and characteristics in Table 2.3 are analyzed by reviewing and comparing studies.

Table 2.3 Online community types and characteristics

Objective	Community type	Innovative contents	Characteristics
New product/service ideas (potential needs/ preferences/ requirements)	Review community	Talk about their experiences of products and services [2]	User with extreme experience
	User forum	Certain information, questions and opinions with common interests [106]	Lead user suggesting potential needs, dissatisfactory points and even feasible solution
Diagnosing product/ service quality (satisfactory or unsatisfactory points)	Brand community	Product or service related information, usage experience [46]	User with brand loyalty
	Social media	Users' attitude on some specific topics, such as consumer product and service	The majority of public users
Exploring feasible solution for existing problems	News and blogs	Expert technology forums and trend review [106]	Expert of users with high interest for specific product and technology

Taking advantage of valuable knowledge beyond big data will become the basis for competition for today's firms [11, 122]. However, one of the challenges concerning the big data consists in understanding how to extract new insights, thus adding valuable knowledge, since they require the use of specific techniques [36].

2.2 Methodological Background

2.2.1 Morphological Analysis (MA)

Morphological analysis is a method for identifying and investigating the total set of possible relationships contained in any given, multi-dimensional problem [135]. The basic idea of morphological analysis is that the system is broken down into several subsystems, through which the system can be described as comprehensively and detailed as possible [125]. This decomposition process can be organized into a morphology matrix, composed of dimensions and shapes [17]. Table 2.4 shows a conventional example of a morphology matrix for a motor cycle. In this example, the motor cycle can be broken down into five dimensions (functions): propulsion, store electrical energy, store gasoline, support driver, and brake. Every dimension includes more than one shape (solution); for example, electrical energy can be stored in lead, NiCd or Li-ion batteries. After developing the morphology matrix, different concepts are created by combining various shapes to form a complete system. As an illustration, the combination of underlined shape represents an environment-friendly concept. This way, new and different forms of a concept that have yet been realized can be thoroughly identified after removing the existing solutions.

Table 2.4: Morphology matrix example for a motorcycle (Ölvander, Lundén, & Gavel [99])

Dimension	Shapes		
	1	2	3
Propulsion	Combustion engine	Electrical motor	Hybrid propulsion
Store electrical energy	Lead battery	NiCd battery	Li-ion battery
Store gasoline	Gasoline tank	No tank	
Support driver	Steel frame	Aluminum frame	Carbon fiber frame
Brake	Disk brake	Drum brake	Regenerative electrical brake

As described, MA is capable of systematically decomposing a subject and then mixing and matching the parts to expand the possibilities for new opportunities of the subject. Such a process is expected to increase the possibility of producing creative ideas.

2.2.2 Textual Case-Based Reasoning (CBR)

Case-based reasoning method is a problem-solving approach that relies on similar prior cases to find a solution [64]. A new problem is solved by finding a similar past case and reusing it to suggest a solution in the new situation [1]. The conventional CBR process is represented in Figure 2.2. A CBR system involves the following three components: (1) a case representation scheme, (2) a similarity metric, and (3) a case-retrieval mechanism [126]. First, the representation scheme considers the case as a model composed of a set of attributes, which characterizes it for a particular application; second, a similarity metric is employed to measure the similarity between the case and those in the database; and finally, a case-retrieval mechanism is utilized to explore the most similar case that is suitable for identifying a solution to the target problem [126].

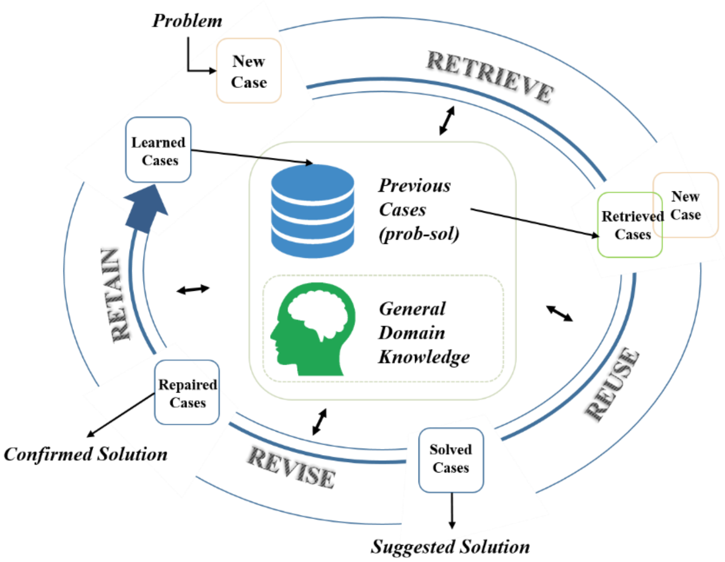

Figure 2.2: The conventional CBR process

Research on CBR in product innovation has emphasized the adaptation phase, which reuses and revises retrieved cases [115]. However, with the advent of massive data-driven CBR, efficient case collection, effective case representation, and accurate case retrieval are becoming increasingly important steps in a CBR system [61, 126].

Textual CBR is a subfield of CBR concerned with research and implementation on case-based reasoners where some or all of the knowledge sources are available in textual format. It aims to use these textual knowledge sources in an automated or semi-automated way for supporting problem-solving through case comparison [124]. Over the years, there has been significant progress addressing the threshold challenge facing textual CBR: how to bring textual knowledge sources to bear in supporting reasoning with cases. Specifically, the research has addressed the following questions: (1) how to map from texts to structured case representations; (2) how to assess similarity between textually represented cases. The research on textual CBR has been growing, however, many applications still depend on handwork in the process of case representation and case development.

2.2.3 Technology Roadmap (TRM)

In general, the term 'roadmap' is defined as a view of a stakeholders as to how to achieve their desired goals [77, 103]. Technology roadmap, therefore, is a kind of roadmap to plan how to develop desired technology under various business environment. For this reason, the generic structure of technology roadmap is a multi-layered and time-base chart, which can show the possible evolutions of technologies, products, and market over time, as shown in Figure 2.3 [77]. Based on this generic structure, many variations have been attempted in order to reflect various kinds of business needs and uncertainties. Phaal et al. [101] suggested several types and formats for technology roadmaps. Lee and Park [77] also proposed several types of technology roadmaps according to the purpose of roadmapping.

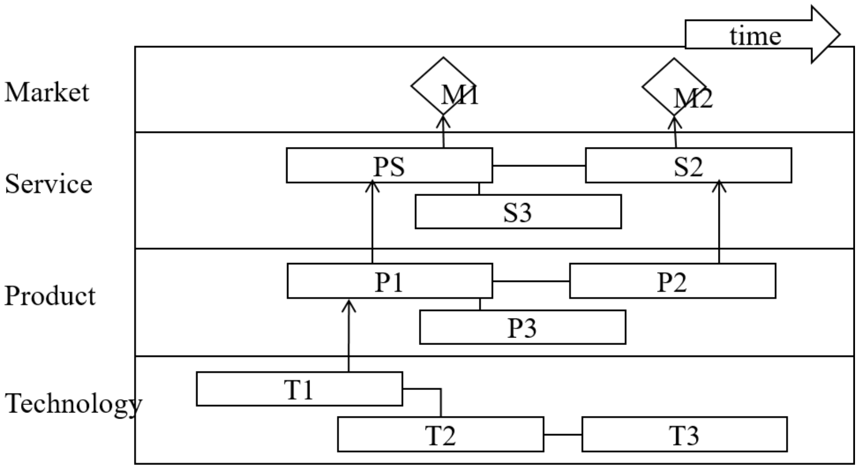

Figure 2.3 Generic structure of technology roadmap

As a means of technology planning, technology roadmaps have been widely used in practice. Technology roadmap, with its graphical structures, can link business strategy to the evolution of product/technologies features [3, 77]. It can be also effectively used for supporting both technology push which aims to look for diverse opportunities, and market pull which aims to develop customer defined products [44, 102]. It is future-oriented strategic planning tool to provide structured approach to linking existing and future technologies, products, services, and market. With its flexibility in the architectural structure and developing processes, technology roadmaps have been widely used in practice [77].

Confronted with recent business changes, several studies have tried new type of technology roadmaps. Geum et al. [44] suggested a product-service integrated roadmap in order to highlight the role of technology in integrating products and services, and provided several types of roadmapping to accommodate different business situations. An et al. [5] suggested how product-service roadmaps can be developed based on revised quality function deployment (QFD). Vishnevskiy et al. [117] suggested an integrated roadmap which can effectively combines market-pull and technology-push, addressing the limitation of overstressing only one-side of innovation.

To support technology roadmapping, many methods and techniques have been extensively used, including both qualitative and quantitative methods. For qualitative methods, a Delphi technique which can properly capture experts' intuitive knowledge and opinion has been widely used. Despite the strength, however, the risk to be biased by some experts is still a void in the literature. Therefore, recent studies have more focused on quantitative approach which emphasizes analytic processes based on data. For this purpose, what have been actively used is patent information. Since the patent has been considered as a proxy for technological innovation, it has been used for analyzing technological trend as well as for identifying potential chances for innovation.

Chapter 3

Planning for System Innovation

3.1 Introduction

In recent years, new advances in information and communication technology (ICT) are bringing products to a new level. Particularly, products generate new types of data, involving rich and spatial distributed identification, as well as historical and sensor data. These products are able to cooperate with other products and even humans through exchanging the data via a network, delivering entirely new services based on monitoring, optimization, remote control, and autonomous adaptation [14]. The intelligence level of service (e.g. personal assistant system) has gradually increased, and services (e.g. U-health care system) that were not feasible in previous years have become possible. In this way, an integration of a smart product and service, described as 'smart service systems', has been recognized as an effective strategy in creating more value for customers [28, 66, 80, 90]. Such phenomenon presents an important perspective to be considered in the context of service development. Considering context and information as core elements in new service development process is, in fact, essential for shaping the smart service system more clearly. However, still under-theorized is the issue of how firms reflect these characteristics in development of new smart service system concepts for achieving and maintaining competitive advantages.

As a tool for generating new ideas, morphological analysis has been widely used in both academia and practice. It is a method to find and analyze all possible solutions from multifaceted and unmeasurable complex problems [7, 59, 125]. The main advantage of this method lies in breaking down a complex system into several dimensions to explore unprecedented and creative ideas [41, 70]. A service system is a complex system in which various components interact. Specifically, a system is configured completely different depending on who uses the product, what sensor and network technology are utilized, and so on. Therefore, a change in a component can cause significant change at a system level. In this sense, morphological analysis has been considered to be a prominent method for deriving new service concepts. New service ideas were presented by performing a morphological analysis with several different techniques, such as conjoint analysis [59] and genetic algorithms [74]. Moreover, Geum, Jeon, and Lee [42] proposed new smart service ideas by conducting a morphological analysis which incorporates technology-push and market-pull perspectives.

Despite its strong possibility for revealing unexpected service concepts, one critical shortcoming of morphological analysis has persisted [42, 59, 74]. The majority of procedures require domain knowledge which can be limited to relatively simple systems and can often be subjective or biased [132]. Specifically, the process of determining the dimensions and shapes is highly dependent on the qualitative judgments by experts [59, 74]. This is the most fundamental step of morphological analysis, as it can have a significant impact on subsequent works and even outcomes [70]. In other words, the more diverse the shapes in the morphology matrix, the more likely it is to combine completely different systems. Furthermore, this participatory process is no longer applicable, especially in the areas where service systems' complexity increases. For these reasons, it is necessary to shift toward a data-driven approach, which could offer more unfamiliar yet novel results objectively.

The potential sources of new service ideas have been enriched, due to the considerable growth of the volume of publicly available data on the Web, where customers and providers can freely communicate their opinions. Accordingly, a vast amount of textual data (service documents) including a detailed description and reviews of the service has been accumulated. However, considering the notion of 'garbage in, garbage out (GIGO)', it is critical to secure comprehensive data that can help analysts explore potential opportunities through a screening process based on novelty and quality. First, from a practical point of view, novel services are recognized as good starting points for exploring service opportunities [95]. This is more apparent in smart service systems because they are characterized by fashion-consciousness, shortened innovation cycles, and excessive competition [72, 108]. Specifically, we define the novelty as an indicator of how different a certain service is from other services. Thus, a novel service in this study means unusual when compared to general services. Second, there is no guarantee that newly launched services will achieve commerciality without customers' satisfaction. This is because customer's satisfaction has great influence on the proliferation of the service [29, 54, 67]. Generally, there exists a great dependency on service quality and customer satisfaction [35, 110]. Hence, it is paramount to view both novelty and quality to be in a crucial relationship as screening criteria.

In response, a scientific approach is necessary to enrich seeding ideas for development of new smart service systems. More specifically, it focuses on the nature of smart service systems to be considered consistently in the overall process. For this purpose, this study conducts a novel method of data-driven morphological analysis considering novelty and quality from service documents. It comprises of two stages: screening the existing smart service systems with high innovative potential and generating new smart service systems. For the first stage, the local outlier factor (LOF) is applied to identify novel services and sentiment analysis is conducted to measure the quality of each service. It is a way to identify what new and trendy smart service systems are being developed. The second stage deals with deriving new smart service systems by generating all possible combinations of a morphology matrix. Particularly, the morphology matrix constructed semi-automatically based on novelty indicators. The proposed approach can be effectively and efficiently used in the stage of concept generation in new service development. In this study, an empirical case study was conducted based on service documents related to smart homes. Generated new smart service system concepts are verified through exploratory investigation on the market. It could indirectly prove the feasibility of the proposed approach if generated concepts had been launched in real services after the time of analysis. Furthermore, it is a pioneering approach in that smart service systems are created with the help of our own developed software system utilizing the vast amount of knowledge from service documents.

3.2 Data-driven Morphological Analysis

Although morphological analysis has been employed in service concept generation for fostering creativity [42, 74], previous studies have mainly focused on the application of morphological analysis without considering the effective construction of morphology matrix [41]. In most studies, construction of a morphology matrix is simply conducted by some experts, showing significant subjectivity and lacking creativity. Moreover, a concrete decomposition cannot be achieved solely by a handful of experts as the product, services, and technologies that make up recent service systems become more diverse. To address this issue, there have been some studies to build the morphology matrix using a quantitative approach. Lee, Seol, and Park [73] proposed a new IT-based service concept generation method using patent analysis, and Kim et al. [59] used information from business method (BM) patents to create technology-based services. However, the application of such data is deemed unsuitable of the following reason. Patents are excellent proxy measures for technology since there are abundant technological terms and descriptions in the patent. On the other hand, they cannot be directly applied to service concept generation in that there is a lack of information on products and services corresponding to application aspects of a certain technology. In fact, this is obviously critical considering that products and services are recognized as a trigger for the development of new smart service systems [96].

In this situation, mobile app service documents can be a great alternative to patent data as there is a lot of information including a detailed description and reviews of the service in various aspects. A channel for users to control and utilize smart service systems is mainly a mobile app service platform. In reality, myriads of smart service systems are being registered in the mobile app market every day.

Each of these apps offers different services incorporating different product and technology features. Furthermore, an extensive volume of mobile app service documents has been accumulated on the Web. The characteristics of the patent document and service description document can be summarized as shown in Table 3.1. However, despite its potential utility, such information is merely expressed in an unstructured textual format, and the massive amount of documents means that, inevitably, the manual perusal of such documents by domain experts is unrealistic. Therefore, it necessitates the development of a scientific approach to extract useful information efficiently, which can be a good source of morphology matrix.

Table 3.1: Comparison of patent document and service description document

	Patent document	Service description document
Type of contents	Large volume of explanations limited to a certain technology with a specific purpose	Short and concise explanations of the key features of a certain service system
Direction of contents	In terms of 'how', it describes the functional process of technologies	In terms of 'what', it describes functions provided by the service and its utility
Diversity of keyword	Containing abundant technical keywords focusing on its functionality	Including diverse keywords related to products and services as well as technologies

3.3 Novelty and Quality in Smart Service System

With the popularity of smartphones and mobile devices, mobile app service markets have drawn much more attention to researchers. According to previous studies [72, 75], mobile app service markets have the following characteristics: fashion-consciousness, shortened innovation cycles, and excessive competition.

High-ranking apps are more likely to attract attention from potential users because the rankings reflect the interests of other users [86]. The number of downloads is increased due to a high level of awareness of a service, and such peer influence presents a high interest to the trend. However, there are few unchallenged leaders in the mobile app service market, and rankings change from time to time. Particularly, a shift toward the newly launched service, which entirely disrupts the existing service, occurs frequently. In this situation, novel services can serve as a good starting point for discovering new business opportunities [95]. Among various novelty detection algorithms, a local outlier factor (LOF) has been used in a lot of research to detect novel keywords or documents from a corpus of text documents [60, 72]. As a density-based novelty detection algorithm, LOF has the advantage of being able to detect local outliers relative to the density of their local neighborhoods rather than to a global view of the data [16]. It outperforms distribution-based, distance-based, and depth-based algorithms when the underlying distribution is unknown, the data volume is large, and patterns of data are incoherent. The LOF of an object is measured by the ratio of the average density of its surrounding objects to its own local density. Generally, the procedure of LOF calculation consists of four steps [16], as follows. First, for each object p in a dataset D, the $k-distance(p)$ is calculated as the Euclidean distance between p and its kth-nearest neighbors, where k is called parameter $MinPts$ defined by analyzer. It restricts the number of nearest neighbors or the minimum size of cluster. Thus, $k-distance(p)$ represents the distance to the farthest end of a neighborhood. Second, for each object q, the reachability distance to p, $reach-dist_k(\mathrm{p,q})$, is derived via $max\{d(p,q), k-distance(p)\}$, where $d(p,q)$ is the Euclidean distance between p and q. If object q is the kth nearest neighbor of p, the two elements of the max clause are the same; if it is inside the cluster, $k-distance(p)$ is chosen, while $d(p,q)$ becomes larger if it is outside the cluster. However, the reachability distance of such objects is always computed as the distance to the farthest end of a neighborhood since we only consider the object q, which is within the kth neighbor. Third, when $N_k(p)$ is defined as the set of $p's$ k-nearest neighbors, the local reachability density, $lrd_k(p)$, is computed as $\frac{k}{\sum_{q \in N_k(p)} reach-dist_k(\mathrm{p,q})}$. The local reachability density indicates how locally dense the neighborhood of p is and thus relevant to a frequent, normal pattern. If the reachability distance is small and the neighborhood can easily reach, such neighborhood is locally concentrated. Finally, the LOF of p with respect to k surrounding objects is derived as $\frac{1}{k}\sum_{q \in N_k(p)} \frac{lrd_k(q)}{lrd_k(p)}$. As a result, if an object corresponds to one of the frequent patterns, its density is similar to those of the neighbors, so the LOF approaches one. Otherwise, the LOF is greater than one and increases as the object is located farther from the normal patterns, since its density is relatively lower than that of normal patterns.

In terms of excessive competition, service providers compete more directly with other developers, since the mobile app market offers a greater range of flexibility in versioning strategies within a category [75]. Once a new service is launched and starts to increase market share, services delivering the same functions have emerged simultaneously in a slightly modified form. In such competitive market environment, the word of mouth effect via review from customers who have purchased and experienced the service has a significant impact on service adoption [62, 75, 84]. Accordingly, research has been conducted to enhance and maintain high levels of customer satisfaction by monitoring and listening to customer reviews. According to the previous study [55], there is a discrepancy between textual evaluation (review contents) and scoring evaluation (review rating scores). Thus, it is a suitable approach to select review contents to derive quality scores rather than just adopting review rating scores. Furthermore, a number of studies have reported that user satisfaction is highly affected by the quality of service, which is a critical factor that encourages users to choose among similar service alternatives. In this sense, customer review contents have been recognized as a useful means of evaluating the service quality by using sentiment analysis [31, 58]. Sentiment analysis is a method for identifying how emotions are expressed in texts and whether such expressions include positive or negative opinions toward a certain product or service [98]. One of the simple techniques begins with sentiment expressions with regard to a given object, then classified into a lexicon of positive and negative based on predefined sentiment dictionaries. As a result, a sentiment score is assigned to a given object through computing quantitative value (positive or negative) of a piece of text [49]. This entails the transformation of textual information into a single number of a certain level of customer satisfaction. However, since lexicons are based on unigrams (i.e. single words), it does not take into account qualifiers before a word, such as in 'not great' or 'no chance'. For many kinds of text in customer reviews, there are a lot of expressions including negative qualifiers. Hence, we utilize a bigram, which is a sequence of two adjacent elements from a string, as a unit of sentiment analysis.

3.4 Proposed Approach

This section examines the overall process, giving a brief explanation of each stage at the same time. The core idea of this paper is the application of novelty detection technique to 'service description database' and sentiment analysis to 'service review database' in order to leverage the textual data into morphological analysis. To this end, a schematic overview of our approach is depicted in Figure 3.1.

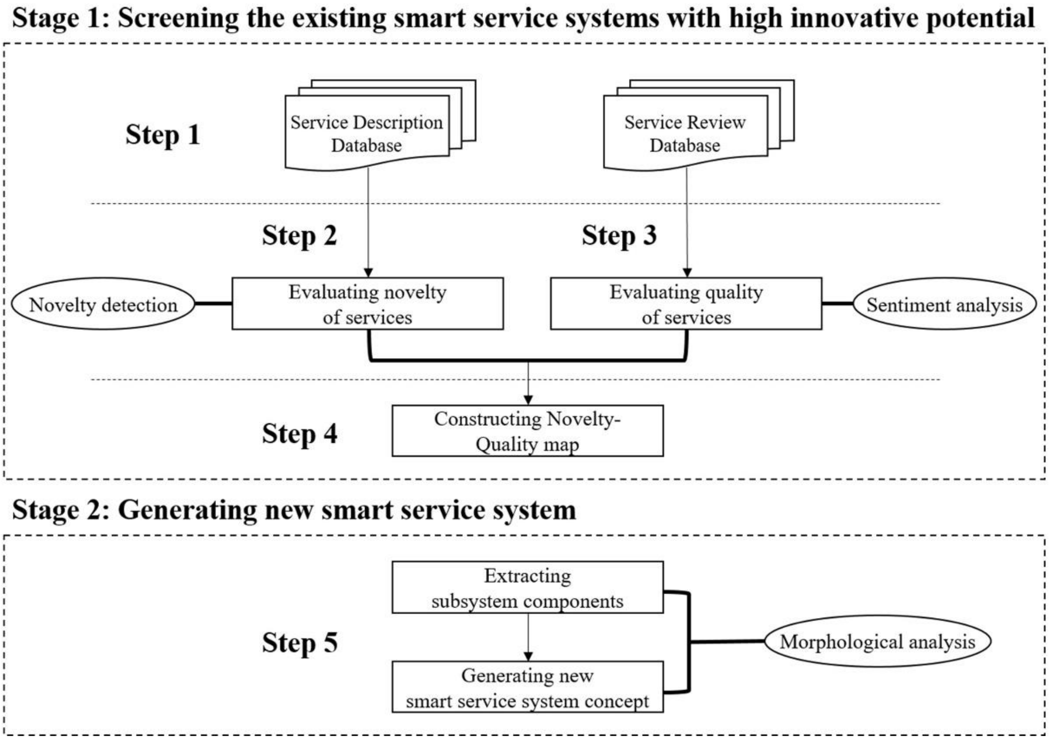

Figure 3.1 A schematic overview of proposed approach

3.4.1 Step 1: Collecting Data

There are websites and online communities with numerous participants where a description and reviews of smart service systems are shared. Once the target service area is determined, documents are collected based on relevant search keywords from the Web. Generally, they consist of unstructured text containing superfluous information such as advertisements to be eliminated. Therefore, text mining is applied to process their natural languages. Keywords are filtered in terms of stop-word removal (e.g., a, the, of, etc.) and stemming.

However, one key thing to note is that pre-processed documents are still not quite fully suitable for applying novelty detection technique and sentiment analysis. Generally, all documents must be transformed into numeric vectors, structured and computer-understandable form according to algorithms. The details are illustrated in further sections.

3.4.2 Step 2: Evaluating Novelty of Services

In this step, the description is employed as an input of the LOF to calculate novelty of services because it can fully explain the differences among existing services [72]. In order to measure the density or distance, the service description documents should be transformed into a term-document matrix (TDM) using n-dimensions of keywords. It is a general method of handling large amounts of unstructured text as analyzable units to extract information [13, 85]. The typical formulation of TDM is described as a two-dimensional matrix as below, whose rows correspond to the terms and columns correspond to the documents, so each $entry\ (i,j)$ represents the frequency of term i in document j. A point of view on the matrix is that each column vector of TDM represents a document with its associated keyword frequencies.

$$TDM = \begin{array}{c|cccc|} & d_1 & d_2 & \dots & d_m \\ t_1 & tf_{11} & tf_{12} & \dots & tf_{1m} \\ t_2 & tf_{21} & tf_{22} & \dots & tf_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ t_n & tf_{n1} & tf_{n2} & \dots & tf_{nm} \end{array}$$

When applied to this research, hth service description document, d_h (where h=1,...,m) is defined as an object of LOF. Since novelty is the infrequency of a service description document within the overall service description database, the LOF for every document is measured as the degree which it is far from the normal patterns. Specifically, the more it includes novel keywords in a service description document, the higher the LOF is calculated.

Novel service description documents are finally identified via the procedure of LOF calculation explained in section 2, after the *MinPts* is defined as *k*. In our approach, the *MinPts* is considered the number of services within local clusters, which are explored to identify novel outliers. Thus, the value of *k* was flexibly changed and determined subjectively according to the context of new service development. For example, using a high value of *k* can identify meaningful novel objects by including more adjacent objects to compare [60, 72]. Although the value of *k* can be determined conveniently by domain experts' judgments, we get an approximate value, applying topic modeling in order to determine the value of *k* more objectively. From the viewpoint of topic modeling, the value of *k* means how many documents are grouped in one topic on average. To derive this information, we utilized the function ‘FindTopicsNumber’ in R packages ‘ldatuning’. Through this, the most preferable number of topics (services) are calculated. Then, based on the given number of topics, we can compare the results of topic modeling to see if a list of topics containing a set of associated keywords represents different service. The final number of nearest neighbors (k) is determined by referring to the average number of documents assigned in each topic. After that, LOF values are computed for each object; the larger the LOF value, the more likely is the object local outlier. If the value is less than or equal to one, such a service document is deemed to be in the frequent or similar-to-frequent pattern and is thus excluded from the set of local outliers [60]. That is, the service descriptions described by the same set of keywords mean frequent pattern. To implement this, we applied the function 'lofactor' in R packages 'DMwR'. The overall process of calculating novelty score is depicted in Figure 3.2.

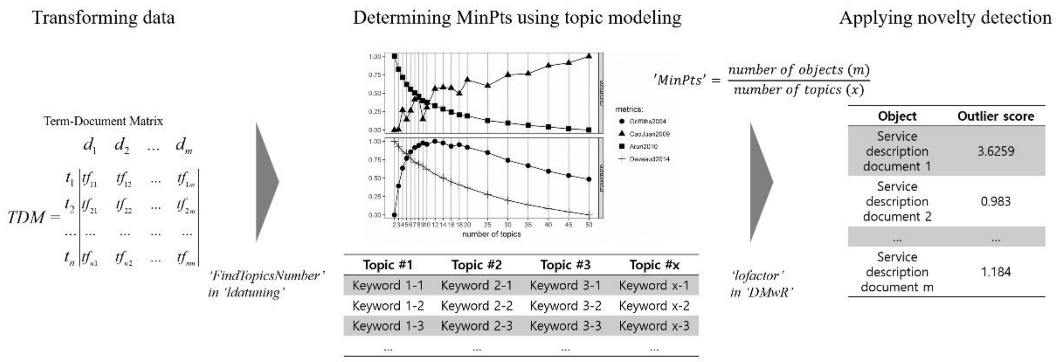

Figure 3.2 Overall process of calculating novelty score

3.4.3 Step 3: Evaluating Quality of Services

In order to assess service quality, lexicon-based sentiment analysis is conducted to determine the polarity of a certain customer review by detecting the occurrence of pre-defined positive/negative keywords. At first, sentiment dictionary is compiled with reference to 'AFINN' [50], which is a list of English words assigned scores for positive/negative sentiment. There are 2,477 words and phrases with an integer between minus five (negative) and plus five (positive).

After that, as was pointed out in section 2, review documents are transformed into unigrams and bigrams to identify sentiment correctly. If the character in the previous sequence contains a negation, the score of the following sentiment word in bigram is applied in reverse. For example, the sentiment score for 'not accepted' is '-1', which is the opposite value corresponding to the sentiment score of 'accepted'. The detailed process of deriving the sentiment score for each review document is illustrated in Figure 3.3.

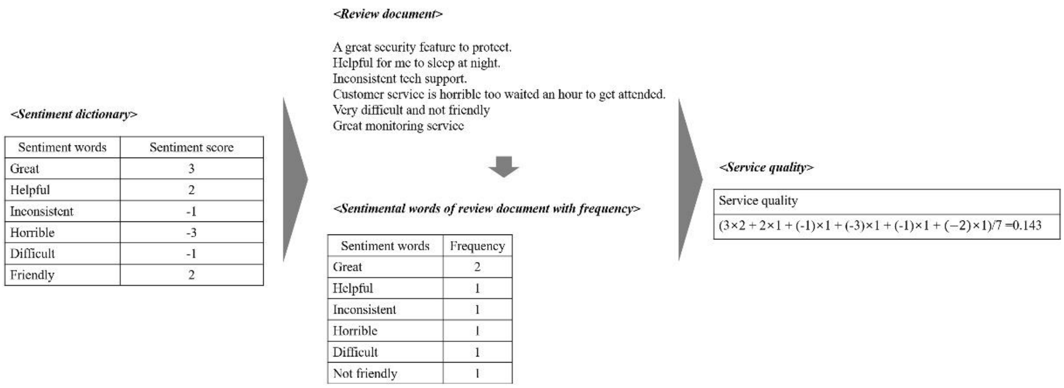

Figure 3.3 An example of lexicon-based sentiment analysis

3.4.4 Step 4: Constructing Novelty-Quality Map

A Novelty-Quality map is constructed to select the services that will be the subject of the morphological analysis, combining both the novelty and quality scores measured through step 2 and step 3. A Novelty-Quality map uses the values of novelty as the horizontal axis and the values of quality as the vertical axis. For relative comparison among services, the center point of the coordinates is the average of the novelty and the quality scores. Consequently, each service can be classified into one of the four areas of the map as shown in Figure 3.4.

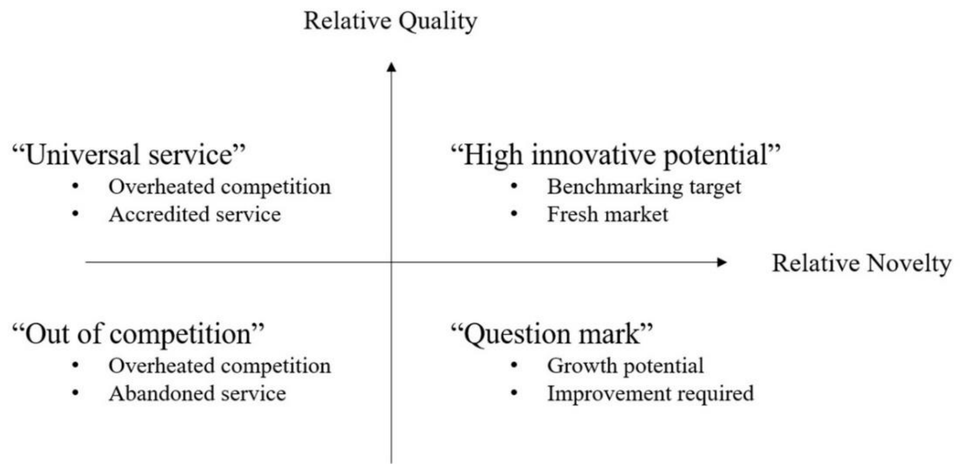

Figure 3.4 The four distinct quadrants of Novelty-Quality map

The interpretations corresponding to the four quadrants are described in Table 3.2. Although appropriate reference services depend on the context of the enterprise, it might be new services with higher quality, which can lead users to purchase and use it. Therefore, this study selects services plotted in the first quadrant for in-depth analysis (i.e. morphological analysis) to generate new smart service system concepts.

Table 3.2 Interpretation of each quadrant of the Novelty-Quality map

Quadrant #	Relative novelty score	Relative quality score	Description
Quadrant 1	High	High	These novel services are likely to have high commerciality since users perceive their service quality as pretty nice. Therefore, they can be a good benchmarking target or reference for new service development.
Quadrant 2	Low	High	There are many similar services in the market. Thus, competition is likely to be very intense. Moreover, as the perceived quality of such services has already been raised, these services are judged to be a regular service.
Quadrant 3	Low	Low	It means a service that does not satisfy users in terms of quality compared to similar type of competing services. Therefore, it can be identified as a service that is out of competition.
Quadrant 4	High	Low	Although, these are new services, they have low perceived quality due to various factors such as technical limit, U/I incompatibility, and price. Thus, it is highly likely to have good commerciality if the users' actual complaints are resolved.

3.4.5 Step 5: Generating New Smart Service System Design

Prior to performing the morphological analysis, the following questions need to be addressed: what aspects of smart service systems should be considered? Although there is no agreed definition of smart service systems, it is reasonable to define them as 'a service system capable of learning, dynamic adaption, and decision-making based upon data received, transmitted, and/or processed to improve its response to a future situation' [94]. It means that introduction of smart devices and data analytics transforms service systems framed as 'a configuration of people, technologies, and other resources that interact with other service systems to create a mutual value smarter than ever [14, 92]. Thus, incorporating product, people, context, technology, service, and information as core elements in generating new smart service system concepts is in fact essential.

Due to this consideration, the dimensions of morphology matrix are determined as product, people, context, technology, service, and information. However, what is important in generating new smart service system concepts by integrating shape of each dimension, is novelty of the shape. This is because the combination of new shapes can lead to a completely different concept of a smart service system. Thus, keywords are extracted from the service description documents located in the first quadrant of the Novelty-Quality map. Furthermore, the LOF value of a keyword is calculated by applying the novelty detection technique in the transposed matrix of TDM, of which column vectors represent a keyword with its appearance in service description documents. In particular, the lower LOF value is calculated as the keyword appears in multiple service description documents. Then, the extracted keywords along with the novelty score are arranged for each dimension of the morphology matrix by domain experts.

Based on the various shapes within the six dimensions, the process of creating a new combination starts with selecting the target product. At first, firms should find a viable product based on current development capabilities including resources, technology, and funds. Once the target product is determined, it is crucial to explore what services to be combined with the product. For this, the following questions should be considered by practitioners: (1) who uses the smart product and under what circumstance, (2) what kind of service can be provided through integration of technology and information associated with the product. It is possible to figure out answers to both questions exemplified in Table 3.3 via generating scenarios with a combination of shapes from the morphology matrix above. Lastly, reasonable and innovative ideas are selected by experts. The overall process of the morphological analysis in this study is shown in Figure 3.5.

Table 3.3 Shape-based scenario generation for exploring service context and service solution

Question	Related dimension	Shape
who uses the smart window under what circumstance? (service context)	People	Keyword X
	Context	Keyword Y
what kind of service can be provided through integration of technology and information? (service solution)	Service	Keyword B
	Technology	Keyword C
	Information	Keyword Z

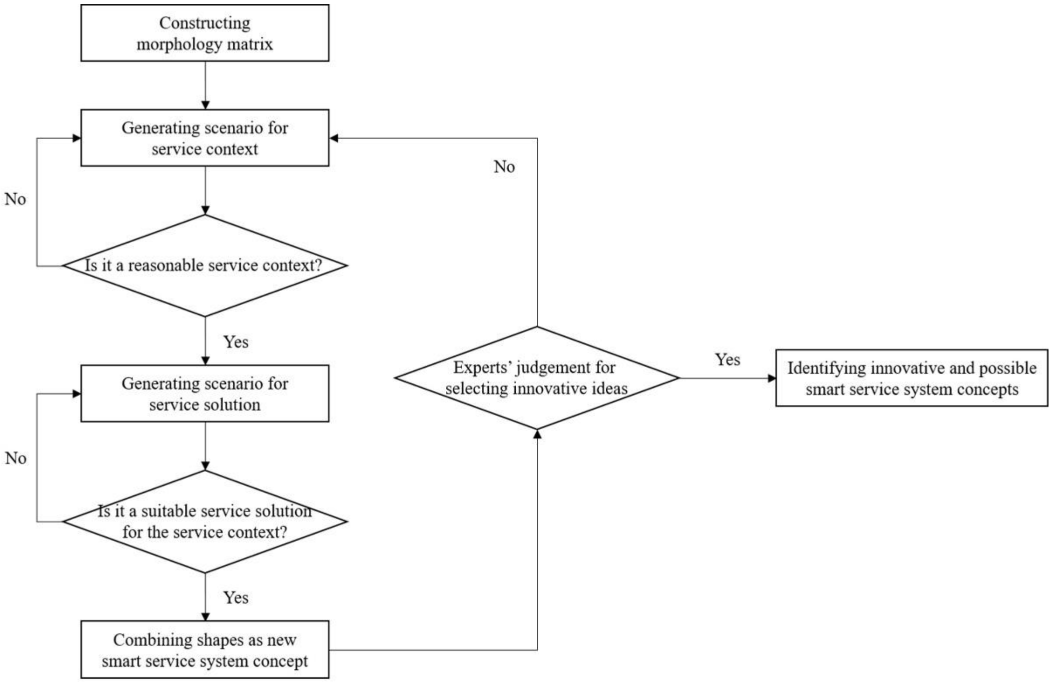

Figure 3.5 The procedure of creating innovative and possible combinations

3.5 Illustrative Case Study: Smart Home

3.5.1 Background

In order to illustrate the applicability of the proposed approach, a case study of 'smart home' was conducted. Based on the rapidly diffused infrastructure of mobile network environments, the demand for smart home services and home security products has increased exponentially [9]. Accordingly, various companies are simultaneously attempting to combine home appliances with new services, creating a highly competitive environment. Typical examples are the remote controls of washing machines, refrigerators, and even utilities such as electricity and water supply with networking capabilities. Therefore, it is very important to have a quick scan of existing smart home services, as it can lead to the discovery of new business opportunities.

3.5.2 Process and Results

We base this case study on the service description documents and review documents that are offered by the App Store and Google Play. They are the most popular open market platform where developers register their services to release and customers share opinions based on their experiences on those services. In our approach, the information on the unique page for each service is divided into the service description and review document. The service description document contains mainly the service description and other useful information such as the price, developer, and evaluation ratings. In the review document, diverse opinions from users are included.

We search services relevant to smart home using query keywords such as 'smart home', 'smart home appliance', and 'smart home service'. However, the number of retrieved services is so large that we are not able to collect all of them manually. For this reason, a Java-based software system was developed to download both the service description documents and review documents automatically. As a result, we collected a total of 1028 services with the reference period from 2010 to 2016. Relatively new services contained a small number of user comments. In this situation, quality scores can be overestimated or underestimated by a few comments when the number of user comments is too small. Thus, we finally selected 513 services in English with the size of the service description being greater than 2KB, and at least 20 user comments.

Before moving to the next step, we applied Natural Language ToolKits (NLTK), which is widely used in performing the Natural Language Process (NLP) for collected textual data. This means that pre-processing was conducted to remove the stop words such as articles, prepositions, and conjunctions. Moreover, the collected documents were parsed to separate out the parts that could be distinguished in terms of information.

The novelty score of service documents was measured via LOF calculation, as described in Section 3. At first, the pre-processed service description documents were converted to TDM to be used as an input of the LOF. With the aid of the Java program for text mining developed by the authors, the TDM consisting of 1,223 keywords with 513 documents was developed as can be seen below in Table 3.4. Since the length of a document could affect the occurrence of its keywords, the normalized frequency of keywords with respect to the length of a document was used as an adjusted value.

Table 3.4 Part of the term-document matrix (TDM)

	Service _id_1	Service _id_2	Service _id_3	...	Service _id_513
TV	4	0	1	...	0
switch	0	3	4	...	2
bulb	0	0	12	...	0
control	6	1	5	...	0
location	2	0	0	...	7
...	...	...	...	...	...
Bluetooth	1	0	2	...	0

Based on the TDM, topic modeling was performed to determine the MinPts. The most reasonable number of topics were selected as 26, therefore, the number of its nearest neighbors (MinPts) is determined as 20. After that, the result of LOF calculations on the description document of each service is derived as shown in Table 3.5. Only 165 of 513 documents have a novelty score; the rest of them have an infinite LOF value because their local reachability densities are at zero.

Table 3.5 Parts of assessing novelty score of services

Service description document id	Novelty score
Service_id_28	1.58
Service _id_50	1.29
Service _id_52	1.09

Service _id_53	1.23
...	...
Service _id_503	1.05
Average	1.17
Standard deviation	0.2167

For evaluating the quality of apps, sentiment words were extracted from collected service review documents based on the words expressed in verb phrases, adjective phrases, and adverbial phrases. Through this, all the service review documents were transformed into sentimental words vector including bigrams and unigrams. With reference to the sentiment dictionary, 'AFINN', the service quality of each review document is calculated as described in section 3. The result is shown in Table 3.6. The subject of quality assessment is limited to the services not having an infinite value of the novelty score. It means that a document that does not contain any keyword derived in the term-document matrix has an infinite value. Thus, further analysis of the term-document matrix becomes meaningless.

Table 3.6 Parts of assessing quality scores for each service

Service description document id	Quality score
App_id_28	2.14
App_id_50	2.85
App_id_52	0.88
App_id_53	1.67

...	...
App_id_503	-0.65
Average	1.12
Standard deviation	1.04

Based on the evaluation of novelty and quality, the Novelty-Quality map was constructed to identify services which have high innovative potential. The simple statistics of the novelty and quality scores are as follows: The maximum value of the novelty score was 2, the minimum value was 0.96, and the average value was 1.17. For the quality score, the maximum value was 3.28, the minimum value was -2.45, and the average value was 1.12. Therefore, the origin of the coordinate axes is (1.17, 1.12). The Novelty-Quality map of 165 services which is based on this grid is shown in Figure 3.6.

Circles in the Novelty-Quality map represent a single service and can be distinguished by 'service_id'; 30 services in the first quadrant, 56 services in the second quadrant, 47 services in the third quadrant, and 32 services in the fourth quadrant. The characteristics of each quadrant in terms of the market perspective are described in section 3.

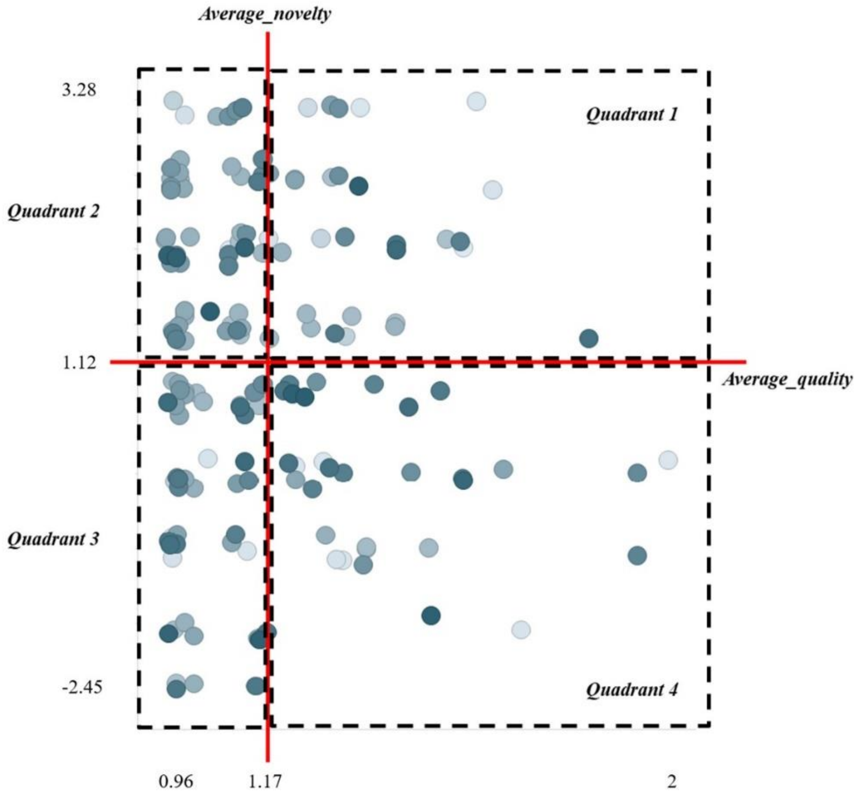

Figure 3.6 Result of constructing Novelty-Quality map

As was pointed out in the detailed process of this step, we performed a morphological analysis on the services which have high scores in both novelty and quality assessments. At first, 202 keywords were extracted from the service description document. At the same time, the LOF value of a keyword was calculated based on the transposed matrix of TDM, used to identify novel services. Then, 202 keywords were assigned as a shape to the product, people, context, service, technology, and information dimensions along with a novelty score by a group of experts as exemplified in Table 3.7. The group comprised three experts having experience in developing mobile app services from three years to seven years and five experts having experience in service management from two years to five years.

Table 3.7 Parts of assigning dimension of each keyword with novelty score

Shape (keyword)	Dimension	Novelty score
forecast	Service	5.5045
breast	Information	6.3581
pet	People	8.7188
iris	Information	4.9979
gyroscope	Technology	3.8120
assistant	Service	4.9139
tracking	Service	4.8106
thermostat	Product	4.7164
virtual reality	Technology	5.1078
...	...	...
hospital	Context	3.3263
customize	Service	3.0175
speaker	Product	2.7302
pressure	Information	2.4184

Through this, the morphology matrix for developing the new smart service system concepts is organized in the following tables. In particular, the keywords assigned to each dimension in Table 3.8 are listed in descending order of their novelty score.

Table 3.8 The morphology matrix for developing new smart service system concepts

Dimension	Shapes (keywords)
Product (37)	pesticide, curtain, sprinkler, fruit, valve, water, vegetable, bulb, heater, window, food, crop, drain, medicine, furniture, thermostat, cctv, smartcam, projector, meal, speaker, car, door, washer, refrigerator, bed, air conditioner, robot, vehicle, switch, watch, book, clock, button, tv, camera, controller
People (19)	pet, dog, cat, neighborhood, family, friend, student, teacher, educator, parent, baby, mother, father, kid, doctor, patient, child, boy, girl
Context (29)	landmark, church, farm, museum, vacation, travel, street, park, apartment, kitchen, garden, garage, subway, road, shop, restaurant, hospital, market, theater, store, inventory, emergency, livingroom, restroom, school, bank, room, office, floor
Service (47)	discover, forecast, estimate, optimize, automate, analyze, view, report, suggest, drive, pay, assistant, support, tracking, purchase, deliver, diagnosis, save, search, buy, monitoring, control, manage, maintenance, provide, fitness, sleep, reduce, customize, alert, recommend, scan, choice, supply, check, warning, shopping, music, navigation, photo, alarm, read, measure, help, call, email, access
Technology (21)	AI, privacy, security, virtual reality, 5G, gyroscope, rotation vector, audio, recording, microphone, ZigBee, gravity, Bluetooth, video, GPS, led, QR, sensor, usb, wifi, battery

Information (49)	breast, iris, tone, skin, ingredient, fingertip, image, finger, voice, face, brightness, mood, gesture, picture, electricity, geography, humidity, heat, pressure, quantity, health, recipe, motion, position, style, status, energy, lighting, quality, temperature, rain, climate, wind, traffic, air, direction, time, color, size, sound, speed, volume, price, map, location, distance, weight, length, weather

For a window manufacturer, let's imagine a situation that explores a smart window idea. In determining the shapes of each dimension, the following questions should be considered by practitioners: (1) who uses the smart window and under what circumstance, (2) what kind of service can be provided through integration of technology and information associated with the smart window. Through this, there can be an innovative idea based on the combination of 'window-student-museum-view-virtual reality-picture', as shown in Table 3.9.

Table 3.9 An example of shape-based scenario generation

Question	Generated scenario
who uses the smart window under what circumstance?	Students who cannot go to museums due to physical limitations want to experience the collections through the five senses
what kind of service can be provided through integration of technology and information?	Normally, it plays the role of showing the same view as the ordinary glass. Various collections including pictures and statues of museums based on virtual reality or augmented reality are displayed through the touch panel.

3.6 Discussion

3.6.1 The Value of Data Source for Designing Smart Service System

This study contributes to the field in that it presents a new data source for morphological analysis by focusing on the development of a new smart service system. As shown in the case study, service description documents contain general but various keywords which are difficult to think of with only a handful of experts. Moreover, the difference from the patent document used as a source in previous studies can be confirmed by the following tables. Table 3.10 shows the extracted keywords from a single patent and service description document with a similar file size. The numbers in parentheses indicate the number of keywords derived from each document. As mentioned in section 2, the service description document contains specific information, which is neither comprehensive nor abstract, about each element that constitutes a smart service system.

Table 3.10 Comparing derived keywords from single patent document and service description document

	Patent document	Service description document
Illustrative case (title)	Dynamic control of smart home using wearable device	Niko Home Control
File size	4 KB	4 KB
Product-related keywords (6/12)	appliance, electronic device, bracelet, miniaturization, necklace, watch,	control, automatic, gas, shutter, water, blinds, button, cabinet, display, wheel, thermostat, window
People-related keywords (1/3))	user	parent, user, pet
Context-related keywords (0/5)		room, home, ventilation, installation, vacation
Service-related keywords (11/13)	control, automation, receive, determine, communication, identify, adjust, analyze, monitor, connect, sync	detection, monitor, search, manage, adjust, find, change, connect, select, check, indicate, lighting, heating
Technology-related keywords (8/11)	sensor, Bluetooth, embodiment, parameter, programmed, network, power, microprocessor,	privacy, Wi-Fi, rotation, panel, IP, gateway, screen, network, Bluetooth, video, audio
Information-related keywords (5/10)	data, health, body, information, location,	preference, data, energy, location, electricity, information, status, consumption, volume, motion

If there is a corpus of documents rather than a single document, will there be different results? The keywords derived from each corpus of 10 documents with similar file sizes have been organized in Table 3.11. Keywords derived from patent corpus clearly reveal that the results are biased to the technological terms. Above all, there is a marked difference in the abundance of keywords related to products and information. That is, the possibility of the newness of the shapes within the morphology matrix is increased, reflecting the characteristics of the smart service systems. Considering the importance of products and information as a trigger for the development of a new smart service system, the usefulness of service description documents is increasingly high. However, technological features are also an important factor to consider. This is because the types and amounts of information that can be obtained depend on the technologies related to sensors, networks, and storage. In this situation, what is important is that technical terms, which are not specifically addressed in the service description documents, are well described in the patent documents. Thus, it is necessary to secure a method that incorporates both patent and service description documents.

Table 3.11 Comparing derived keywords from 10 patent documents and service description documents

	10 patent documents	10 service description documents
File size	24 KB	28 KB
Product-related keywords (7/14)	device, vehicle, handle, storage, light, computer, CD-ROM	device, optical, storage, light, DVD, led, switch, tv, blind, curtain, appliance,

		projector, switch, water
People-related keywords (3/6)	driver, user, owner	kids, users, parents, children, family, friend
Context-related keywords (3/5)	house, home, office	home, inventory, room, floor, showroom
Service-related keywords (15/16)	control, monitor, charge, determine, automation, provide, detection, identify, read, receive, show, support, purchase, share, deliver	control, support, provide, find, heating, automation, call, enjoy, charge, monitor, alarm, buy, check, protect, contact, create
Technology-related keywords (24/17)	network, internet, server, laser, protocol, security, module, wave, propagation, multiport, plug, router, video, battery, display, engine, motor, photodiode, socket, wafer, cable, Wi-Fi, converter, privacy	network, interface, IP, screen, touch, display, gateway, server, module, USB, video, Wi-Fi, camera, privacy, security, beacon, Bluetooth
Information-related keywords (9/16)	signal, time, energy, data, voice, location, size, speed, media	data, action, consumption, status, time, information, location, media, sound, barcode, energy, climate, temperature, weather, brightness, motion

3.6.2 The Role of Novelty-Quality Map

The morphology matrix for development of smart service systems is constructed semi-automatically. We carefully reviewed previous literature and identified key elements that can adequately describe a certain smart service system. In particular, we defined 6 dimensions based on the definition of a smart service system, adopted in both academia and industry. There might be several ways to define dimensions, and this could be selected differently according to the firms' circumstances. For example, when a specific target product is determined, one can divide each dimension into more sub-categories such as parts, size, color, shape and other features of the product. Therefore, the involvement of experts' judgment in the process of determining the dimensions is inevitable.

Instead, we referred to keywords derived automatically from a large amount of service description documents to determine shapes. In this process, data screening through a Novelty-Quality map has the following implications. From a practical perspective, it helps you to quickly scan existing smart service systems by visualizing them as a sort of positioning map. Depending on the location plotted on the map, the innovation direction and strategy may be different. For a service in the third quadrant, it can take a strategy to benchmark similar services in the second quadrant or stop providing the service. From a methodological point of view, it makes shapes of the morphology matrix include more creative and unusual keywords. The novelty score of the service description document becomes higher when the document includes keywords which are not included in other service description documents. Moreover, the quality score refers to the evaluation of a smart service system that includes a novel element, thereby reducing the market uncertainty about considering the novel element as an alternative. The high level of satisfaction with the smart service system, even though it includes a novel element, confirms the commerciality. Thus, utilizing service description documents with a high novelty and quality score, the first quadrant of the Novelty-Quality map, as input data is reasonable for constructing inventive shapes. The results of keyword extraction from the arbitrary service description document plotted in each quadrant are shown in Table 3.12. The keywords derived from the service description documents located in the second and third quadrants with low novelty score are ordinary compared to those of the first and fourth quadrants.

Table 3.12 The comparative keyword extraction results of each quadrant of the Novelty-Quality map

	Quadrant 1 High innovative potential	Quadrant 2 Universal service	Quadrant 3 Out of competition	Quadrant 4 Question mark
Illustrative case	App_id_285	App_id_153	App_id_88	App_id_55
Product	smartcam, device, builtin, camera	watch, device, smartphone	camera, device	vehicle, button, car, door, lighting, thermostat
People	traveler, user	user	children, user	guest, driver, family, kids
Context	home, hotel, leisure, holiday, residence, office, emergency		home	home, house, parking, emergency, garage, road
Service	store, send, support, alert, detect, integrate, analyze, control, email, record	synchronize, detect, email, export, import, protect, access, search, share, SMS, store	capture, control, hear, monitor, photo, record, support	manage, monitor, check, control, protect, alarm, alert, assist, remind, provide, streaming
Technology	video, audio, security, sensor, server, network, router, Wi-Fi	touch id, cloud	video, audio	security, motor, video, GPS, network, router
Information	motion, face, signal, quality, frequency, schedule, picture, sound, time	data, size, password, date	gesture, sound	fee, location, time

3.6.3 Feasibility and Effectiveness of Proposed Approach

Although we conducted the case study for illustration, the feasibility and effectiveness of the proposed approach cannot be strictly verified. Therefore, further analysis is performed to confirm whether a new generated combination has been released after the time of data collection. More specifically, the results of the morphological analysis are compared with real service examples. In terms of feasibility, if the newly generated smart home service concepts are searched in the 'app store' and 'android market', this process can be considered as an experiment for partial verification. Through this, it was found that 10 of the 15 new concepts created by the experts have been taken by App developers and accepted into the market, as shown in Table 3.13. On the other hand, the five new concepts have not been developed. It can be a starting point from which, in terms of effectiveness, it can provide an entirely new opportunity for service creation. In other words, the proposed approach enables the creation of a service concept that includes elements which are difficult to imagine.

Table 3.13 Comparing new smart home service concepts with real services

Derived new smart home service concepts in 2016			Real service examples launched after 2016		
No	Product-People-Context-Service-Technology-Information	Description	Service name	Launch date	Relevancy
1	Sprinkler-Neighborhood-Garden-Forecast-Sensor-Humidity	Adjusting the spray amount of the sprinkler to prevent water waste by forecasting the weather based on the humidity information shared through the networked sensor of the neighboring sprinklers	Blackloud Sprinkler	2017. 7. 24	Customizing watering schedules with weather information and controlling smart sprinkler anywhere, anytime
2	Bulb-Family-Livingroom-Optimize-Sensor-Motion	Optimizing the illumination of a bulb by identifying the behavior of the people surrounding the bulb	Philips Hue	2016. 8. 9	Customizing the color and illumination of a bulb remotely via voice recognition
3	Crop-Family-Farm-Report-Video-Map	Reporting and monitoring whether the crops are growing normally by visualizing the farm as a map through video	R Phoenix farms	2016. 3. 2	Reporting on sowing quality, stand counts, and plant health through drone mapping
4	Smartcam-Baby-Emergency-Monitoring-Video-Position	Monitoring baby in real time for dangerous movements or emergency situations via Smartcam	Dormi – Baby Monitor	2016. 3. 22	Monitoring baby along with video streaming at sleep and activating alarm for parents

					in emergency situation
5	Smartcam-Pet-Emergency-Analyze-Video-Motion	Analyzing health status by monitoring pet behavior via Smartcam in absence	Alfred home security	2016. 3	Monitoring pet behavior and activating instant alerts when something happens
6	Medicine-Patient-Room-Assistant-Recording-Time	Assisting the correct medication through alarm function with displaying the time of dosing in the vial	MediAlarm	2016. 8. 25	Tracking the medication history and helping the patient not forget taking medicine
7	Refrigerator-Parent-Kitchen-Purchase-Camera-Ingredient	Purchasing the ingredients for specific meals after monitoring inventory in the refrigerator	Samsung Family Hub	2017. 3. 3	Helping to purchase consumer goods considering preference of family member through a simple process
8	Curtain-Family-Room-Automate-Sensor-Motion	Adjusting curtain or blind position automatically by detecting motion of people or voice recognition	MySmartBlinds (2.0)	2017. 11. 5	Adjusting blind position automatically by filing absence time at home
9	Smartcam-Family-Apartment-Customize-Security-Iris	Customizing smart home security system with iris, voice, face, and finger for improving security level	Iris by Lowe's	2016. 6. 18	Managing, monitoring, and controlling smart home devices from anywhere through iris security system
10	Window-Family-Room-Suggest-Virtual reality-Style	Suggesting and recommending trendy fashion style by showing a person wearing those clothes through virtual reality	Amazon – Shopping made easy	2017. 9. 27	Supporting users to preview the furniture in their living room before purchase

11	Pesticide-Family-Farm-Tracking-Video-Image	Tracking pests around the farm through video, and autonomously distribute the appropriate amount of pesticide	Relevant services are not developed or released
12	Furniture-Patient-Emergency-Automate-AI-Temperature	Designing furniture to maintain body temperature automatically to improve patient immunity	
13	Water-Kid-Restroom-Reduce-Sensor-Humidity	Wiping the water off automatically to prevent kids from slipping by sensing the humidity on the floor of the restroom	
14	Food-Doctor-Kitchen-Report-Sensor-Quantity	Reporting the amount of meal, salt intake, and etc. to doctor for controlling the patient's diet	
15	Valve-Family-Vacation-Control-Sensor-Speed	Controlling valves to save water, electricity, gas, and etc. via sensors attached to them, when family vacate home	

However, the results of the proposed approach are nothing more than an idea generation that is a very small part of the whole process of developing a new service. In order to realization of the smart service system concepts, firms should consider a lot of issues such as law, regulations, resources, and technical competence. Therefore, it is important that the results should be effectively linked to the full-fledged service development process.

3.7 Conclusion

We suggested a novel data-driven approach to the generation of new smart service system concepts, which consists of two stages. The smart service systems which have high innovative potential are identified at the first stage, based on novelty and quality measures. Novelty detection technique and sentiment analysis were employed here as a tool for dealing with mobile app service documents. Then, a number of new smart service system concepts were derived at the second stage by exploring all possible combinations of a morphology matrix. Furthermore, a case study on 'smart home' has been illustrated to show the feasibility.

The benefits of the proposed approach are as follows. First, from a theoretical perspective, this paper suggests a new and effective way to scan the current market. The existing smart service systems are classified into four categories, which are based on novelty and quality. With a Novelty-Quality map, one can easily identify what new and trendy smart service systems are being developed and released to the public. Second, from a methodological perspective, this paper proposed a quantitative approach using mobile app service documents as a remedy for subjective and expert-based morphology building. Specifically, dimensions are developed based on the characteristics of smart service systems, whereas shapes are derived from service description documents via the combination of textminig techniques. With this morphology matrix, a large number of alternatives are derived in a very short time. Third, from a practical perspective, the proposed approach provides managerial implications for the practitioners who are in charge of the new service development. It enables practitioners to capture diverse and unusual elements that are to be considered as a trigger for the development of new smart service systems. This can lead to new ideas completely different from previously existing ones. Particularly, the use of mobile app service documents provides a practical implication to the use of big data in the innovation process, by facilitating both quick decision-making and plentiful decision alternatives. Furthermore, once the overall process suggested in this study has been established, real-time monitoring is possible only by updating input data. This allows the practitioners to identify market trend through two indicators, novelty and quality, among new services that are pouring out.

Despite these contributions, however, this research is still subject to limitations. First and foremost, as an inherent limitation of morphology analysis, it is unable to combine multiple shapes per dimension. This means that it highly depends on domain experts to derive a service concept, which is packaged together with various services. Thus, specific measures and criteria to identify services that are likely to be bundled into one smart service system would be valuable for future research regarding morphological analysis. Second, services corresponding to the fourth quadrant of the proposed Novelty-Quality map could also be an incremental innovation target. In that case, for the purpose of improving the existing services, it is more important to identify specific complaints about the service. Therefore, it is necessary to analyze reviews to preserve the meaning of the sentence itself, rather than the keyword level. Third, guidelines for selecting meaningful smart service system concepts are required for the development of new service contexts for each and every firm. Since a large number of combinations are generated based on a morphology matrix, appropriate screening methods and procedures should be prepared.

Chapter 4

Planning for Technology Innovation

4.1 Introduction

Recent innovations have shifted towards customer-oriented innovation. With the rise of connected marketplace, customers become smart due to the explosively generated information from the web. They share information on product and technology and their experience in online communities such as website, forums, and blogs [37, 88, 93], thus have been increasing bargaining power based on well-developed knowledge. For this reason, firms to develop new products are trying to collaborate with these innovative users in the value co-creation processes [93]. These innovative users can be considered to be lead users that can foster the innovation process.

According to von Hippel's work [118], lead users can be defined based on the following two characteristics: (1) they generally face some strong needs before the mass customers encounter them (2) they solve their problems themselves [88, 93, 112, 118]. Taken together, the main characteristics of lead users are simple: they are ahead of the market, anticipating the general needs of the mass market and seeking the solution in advance [88, 93]. Due to this singularity, the lead user technique has been widely used in practice in order to help firms overcome the uncertainties of technological development [104].

However, in practice, lead user technique is hard to implement on account of the time, effort, and cost for identifying and maintaining lead user groups. Therefore, recent studies try to focus on utilizing massive online data for understanding lead user groups [6, 93, 112, 114] and identifying problems and solutions [45, 61, 112]. Data-driven approach for replacing lead user techniques has been actively employed, by leveraging ideas from lead user communities [61].

Especially, the use of case-based reasoning (CBR) fits the purpose of utilizing the data-driven lead user technique, because the core essence of lead users is directly related to the problem-solving processes of CBR. Case-based reasoning is a decision support technique to solve new problems based on existing problem-solving of similar cases [51, 87, 97]. As many works noted [61, 93, 118], lead users are innovative customer groups that face the needs and problems prior to the mass customers and find available solutions by themselves. This means both problems and solutions - which can be effectively utilized as important data for CBR - are expressed in the lead user communities. However, even if CBR is very helpful to utilize lead users' opinion in practice, how the contents of the lead user community can be applied to the CBR still remains a void in the literature. Even if thousands of posts are created from online lead user communities, it is difficult to monitoring all posts in the web. This leads to the necessity of organized guidelines for utilizing such fruitful contents generated by lead users [39].

To address the necessities of models for utilizing online data to CBR, several approaches have been employed to provide a way of data-driven CBR [45, 51, 52, 61, 97, 107, 133, 134], by employing textual data to the CBR process. Even if previous studies worth the effort, they have common limitations in application to the innovation of technology-intensive product from the following two aspects. First, from a practical perspective, the components' diversity has increased as the product's technological intensity increases. In addition, it is a challenging issue to fully understand cause and effect chains among the components [63]. Therefore, unexpected potential problems are likely to emerge in the process of solving a certain problem. There are even cases where defining problems associated with a particular function is difficult. This makes the problem statement, the most important first step of the CBR process, obscure. In this situation, rather than starting with a specific problem, the process of searching for an existing or even, potential problems based on the technological structure of the product should be preceded, which can promote a more concrete and proactive decision-making. Second, from a procedural perspective of textual CBR, previous studies have utilized the keywords as inputs (i.e. problems), calculated keyword similarities as a retrieval algorithm, and derived the documents as outputs (i.e. solutions). However, keywords themselves are not enough to state the problem and hard to be interpreted meaningfully. Moreover, the similarity measure between the paired keyword vectors that represent a problem and a case respectively is not a logical process for finding a genuine solution. In particular, what is important in reflecting the characteristics of the lead user in the textual CBR process is that retrieval process should be rational and meaningful to pinpoint the exact problems and find the links between the problem and the solution.

To address these limitations, this study employs a novel approach to the textual CBR, which explores substantial problems from the database, and suggests how solutions can be derived from the problem set. This study is differentiated from previous works by not only preparing the procedure to consider the context of technology-intensive product innovation but also improving the technical issues of the textual CBR process, by leveraging the lead user communities systematically.

The remainder of this paper is organized as follows. The related works section deals with the background of online lead user communities and textual CBR in innovation research, which are the key theme of this paper. Then, the concept and strong points of methods utilized for an elaboration of textual CBR, are described. The research framework then addresses the proposed approach and its detailed procedures. The subsequent case study presents how the proposed approach can be applied. Finally, our discussion and concluding remarks are provided.

4.2 Ideas from Online Lead User Communities

In terms of value co-creation, attempts to involve lead users in innovation processes have persisted [89]. In particular, the rise of 'Web 2.0' has enabled new ways to apply the lead user approach: innovation by members of online communities [30]. The users in these communities tend to openly share their information, knowledge, and even technologies with other members, exchanging their experiences and developing innovative ideas [61, 65]. Open source software is probably the most well-known example, where geographically dispersed individuals collectively develop new software and product innovation [121].

The contribution of these online lead user communities can be further enhanced in the process of developing technology-intensive products. Particularly, in the case of products such as VR (virtual reality) and drone, there can be a wide variety of problems that lead users encounter. In terms of technology diffusion, these problems prevent the spread to the early majority of a market, so that it is buried in a "chasm" for a long time. As discussed in many studies, a scientific approach to solve the discovered problems and identify potential needs preemptively has been regarded as a key strategy to overcome the chasm. Thus, the distinctive characteristics of the lead user and online lead user community are appropriate for this strategic context. That is, a shared post that represents a specific problem or a problem-solving case can serve as a starting point for developing improvements to existing products or entirely new products. Despite these high utilization of the online lead user communities in the innovation process of technology-intensive products, there is a lack of research on the application of these textual contents. It leads to the necessity of an organized approach to reflect the problem-solving context, which is one of the main characteristics of posts in lead user communities.

4.3 Textual Case-Based Reasoning for Problem-Solving

Most of the studies applying textual CBR [45, 61] use natural language processing techniques like text mining to make unstructured data coherent. Case representation and case retrieval, i.e., input and output of CBR, are based on the keyword vector. Thus, the retrieval method was forced to adopt a simple approach of measuring similarity score between keyword vectors. However, this has the following critical limitations. First, loss of information that does not preserve the full meaning of the document or sentence occurs. Even if it consists of the same keywords, there are various situations depending on the context. Moreover, if the similar case is retrieved, the keywords themselves cannot be a genuine solution. A document (case) generally consists of sentences describing either a problem or a solution. Therefore, it is required to distinguish individual sentences within a document into problems or solutions. To do this, a method should be devised to represent sentences within a range that minimizes the loss of information. Second, in the case of technology-intensive products, a problem-solving is usually knowledge-dependent task because their components are highly complex, and thus knowledge support is inevitable to identify adequate solutions. For example, changing the specification of a particular component to solve a problem may entail changes to another associated component. Thus, it is essential that the relationships among the components should be ascertained proactively. In other words, experts' knowledge must be incorporated into the process of textual CBR, i.e. case retrieval, to identify the exact problems and solutions.

4.3.1 SAO Structures and Semantic Analysis with Sentiment Score

This paper focuses on SAO structures in documents, which are the syntactically ordered structure of subject (noun phrase), action (verb phrase) and object (noun phrase). While SAO is not a deep representation of the sentence, it is more expressive than a keyword vector, offering potential improvements in the performance of interpreting meaning. For this reason, SAO analysis has been used primarily in the handling of technological documents [24, 48, 130]. This structure clearly provides a relationship between components that appear in a technological text: subjects and objects may refer to components of a system, and actions may refer to functions performed by a certain component [20, 22, 129]. That is, SAO structures are fundamentally related to the concept of function, which is defined as "the action changing a feature of an object" [105].

An SAO structure of patent documents can be organized in a problem-solution format if the action-object (AO) forms the problem and the subject (S) states the solution. However, posts of the lead user communities have a different nature from the patent document. While the patent documents focus on what the technology provides to solve the problem, in general, the lead user's experience-based posts focus on the description of the problem situation encountered, and suggest the process of their own conceptual solution. Thus, this study utilizes two distinguish types of SAO structures, depending on whether they represent a problem situation or a solution, instead of using an SAO itself as a problem-solution format. Especially, the semantic analysis with sentiment score fits this approach due to the following reasons: from the semantic relationship of the SAO elements and emotional expressions including positive or negative words, it can be inferred whether it represents a problem situation or a solution. Thus, a system is developed in this study to automatically classify a vast amount of SAO structures as Problem SAO (P-SAO) and Solution SAO (S-SAO) in the case representation stage.

4.3.2 Technology Tree and Co-occurrence Analysis

A technology tree is a branching diagram that expresses relationships among product components, technologies, or functions of a technology in a specific technology area [23]. A function-based technology tree depicts the functions of technology applied to products and their interrelatedness. This is the process and outcome of deriving a set of functions needed to meet the performance goals of the product. For this reason, the technology tree has been used as an important decision tool for technology planning, supporting to define the elemental function based on core technologies under specific goals.

Most technology tree diagrams are created by reflecting the opinions of domain experts in a qualitative manner. It means that the process of constructing the technology tree is highly dependent on the experts' knowledge. However, in the case of technology-intensive products, constructing the technology tree has become more difficult as technological development speed is very fast and the system components are becoming diverse and complex. Often, no detailed information and knowledge is available until the technology-intensive products are widely adopted and utilized [68].

In order to address this issue, the co-occurrence analysis is used as a complementary tool in a quantitative manner. The co-occurrence analysis is simply the counting of paired data within a collection unit. The collection unit and the data therein can be varied. In text-mining, the data is occurrence of the keyword, and the collection unit is the particular document or sentence. It refers to semantic proximity in linguistic meaning and is used to find collocation of keywords. Two keywords with frequent co-occurrence are likely to have a deep relationship in a specific context. In the context of the posts describing the experience that users encounter in the middle of using technology-intensive products, it can be interpreted as follows. Frequently emerging keywords related to the main functional keyword (i.e. problem) are likely to be technically related: (1) They are core elements to improve the function or (2) other operational issues that can be caused during the problem-solving process. Thus, they can be hidden elements to be considered together, which were not found in the technology tree constructed by the experts. Through the hybrid approach of technology tree and co-occurrence analysis, we define the 'mediating keywords' set. It is utilized as an extended input to the case retrieval process to broaden the scope of the case search and consequently enrich the problems and solutions.

4.4 Proposed Approach

This paper proposes a revised textual case-based reasoning approach to extend lead user's knowledge for supporting the development of technology-intensive products. In our approach, the nature of lead user community and the context of technology-intensive product innovation are considered. Unlike the conventional CBR for product innovation, where problem statement is clearly defined, the goal of our process is to identify not only the solution for the existing problems but also potential issues including new needs.

For these reasons, the following Table 4.1 shows the differences between the textual CBRs utilized in innovation research and the revised textual CBR proposed in this study. On the procedural side, it starts with a case collection. Due to the nature of the lead user community, a post has several descriptions of problems and solutions. Therefore, it is impossible to map a single post to a single problem or solution. If a post is used as a case itself, it cannot be a logical process to compare with the problem statement. Because, unnecessary terms are included in the comparison with the problem statement when a post is expressed as a keyword vector. To address this issue, SAO structures in a post that minimize loss of information are classified into problem (P-SAO) or solution (S-SAO) based on semantic and sentiment analysis. For next, a main functional keyword at the heart of the problem is selected as an input for constructing mediating keywords in case retrieval stage. These keywords are derived by complementary use of the quantitative approach (i.e. co-occurrence analysis) and the qualitative approach (i.e. technology tree). Finally, all the retrieved problems and solutions including mediating keywords are evaluated through the following two steps: identification of the critical functions and features associated with the main function, followed by judgement whether the solution is feasible with the experts. Detailed descriptions of these processes are discussed in the next section.

Table 4.1 Comparison of previous textual CBR and proposed approach

Process stage	Previous textual CBR	Revised textual CBR (suggested)
Approach	Data-driven approach	
Data source	Teacher's experiences with technology, Appstore [45, 61], China Academy of Building Research [128]	Lead user community
Case collection	Type1: Accumulating cases without distinction between problem and solution Type2: Constructing problem case and solution case separately from different data sources	Classifying problem and solution SAO structures automatically from each case through two-step filtering
Case representation	A keyword vector with a value of tf-idf or frequency	A Subject-Action-Object (SAO) structure with a sentiment score
Case retrieval	Input: a keyword vector as a problem statement Method: calculating similarity score (cosine similarity) between paired keyword vectors Output: a list of keywords or documents with a similarity score	Input: a main functional keyword as a problem statement Method: constructing a mediating keywords (extended input) through hybrid approach of technology tree and co-occurrence analysis and then deriving all the SAO structures containing the mediating keywords Output: all the problems and solutions with assigned label of function

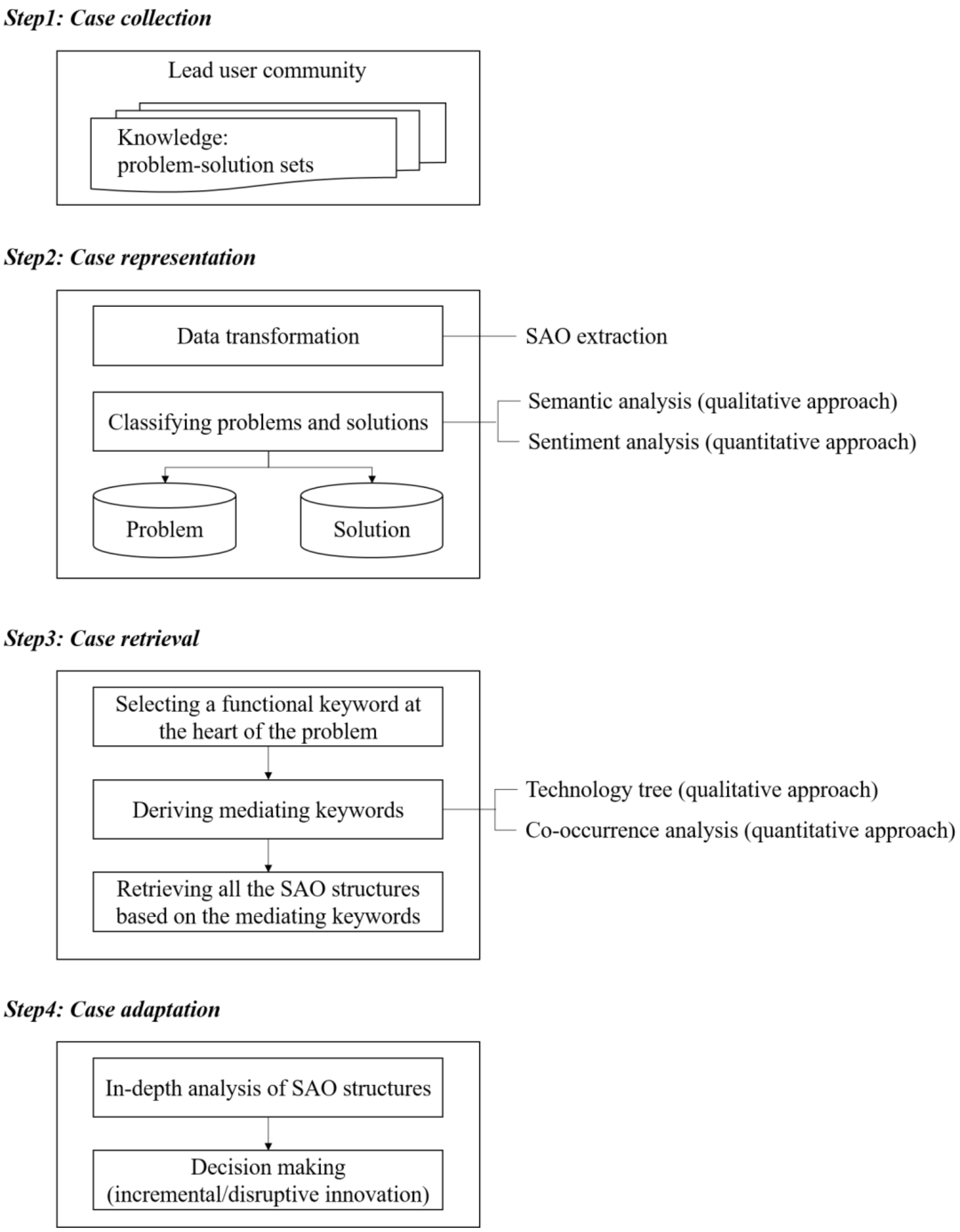

Figure 4.1 Overall process of the proposed approach

4.4.1 Step 1: Case Collection

The main objective of this step is preparing cases of 'knowledge: problems and solutions', which are suitable for data-driven CBR. The most important thing to do is selecting an appropriate lead user community. This is because data-driven approaches depend heavily on the quality of input data. Even more simply, bad inputs will lead to bad outputs, which is so called 'Garbage in, Garbage out (GIGO)'. Therefore, there should be sufficient cases dealing with various issues about the target technology-intensive product in a lead user community.

Since the cases are collected from websites, web-crawlers can be developed and utilized for automatic collection. However, they consist of unstructured text containing superfluous information such as email address, website address, and etc. to be eliminated. Thus it is inevitable to conduct pre-processing such as stemming, stop-word removal (e.g. a, the, of, etc.), and number removal. As a result, each post (case) is collected in the form of a pre-processed document.

4.4.2 Step 2: Case Representation

The case representation stage is a transformation of collected cases to suit the textual CBR format. As discussed, SAO structures are extracted from each document to construct problem and solution database separately. To conduct this, a modified version of the Java-based program called 'ReVerb' is utilized. It takes raw text as input, and outputs (argument1=subject, relation phrase=action, argument2=object) triples. For example, given the sentence "Lead user community is a potential source of innovation," ReVerb extracts the triple (lead user community, is a potential source of, innovation).

After that, we established the criteria for classifying the extracted SAO structures into problem and solution databases. Based on the previous studies and data exploration, we assume that the SAO structures describing a problem or solution have distinctive characteristics summarized in Table 4.2. In general, nouns and adjectives are included in "subject" and "object", on the other hand, verbs and adverbs are mixed within "action" [12]. This structure educates us on the features like 'who (subject) does what action (verb) to whom (object)' [124] and hence, the opinion or action of the subject on the object can be identified. Especially, the verb with adverbs describes the feeling subject has on the object [21]. Moreover, through several posts randomly selected from lead user communities, representative terms describing problem situations and solutions, are derived respectively. A list of synonyms is extracted based on WordNet, which has been widely used in text-mining.

Table 4.2 Distinctive characteristics of SAO structures describing problems and solutions

	Element	Problem description	Solution description
Semantic perspective	verbs in 'action'	Pattern of sentence structure	
		Subject – Synonyms – Function/Feature/Specification Synonyms: need, require, seek, necessitate, want, lack, shortage	Subject – Synonyms – Solution/Function Synonyms: can, could, enable, use, utilize, apply, devise, develop, design, invent, inspire
Emotional perspective	adjectives and adverbs in 'SAO'	Negative adjectives and adverbs	Positive adjectives and adverbs

Based on these characteristics, the semantic analysis, which focuses on the verbs contained in the action, is utilized as the primary filter, then the sentiment analysis for adjectives and adverbs is used as the secondary filter. The SAO structures are classified into problems and solutions depending on whether or not synonyms are included in the action phrase. In the case of not including synonyms, the calculated sentiment score is utilized as a classification index. The analyst should set the cut-off value based on the statistics for the sentiment score of the SAO structures, and then perform the classification process. Note that, the important thing in this step is not to improve the accuracy of the classification. The eventual objective is to eliminate unnecessary sentences in a document and increase the probability of leaving useful sentences to be analyzed.

4.4.3 Step 3: Case Retrieval

After constructing the two databases, a case retrieval process is applied. First of all, a main functional keyword describing the problem situation most precisely, is determined as an input. This paper defines the main functional keyword as 'Root function’. Then, a technology tree is created by an expert group to understand overall technological structure based on the root function as shown in Figure 4.2. Considering the functional relationship, the key functions affecting the ‘Root function’ are placed in the lower layer, and related features are listed sequentially in the substructure of the function layer. As a result, the sub features in the bottom layer of the technology tree are selected as an experts-based mediating keywords list.

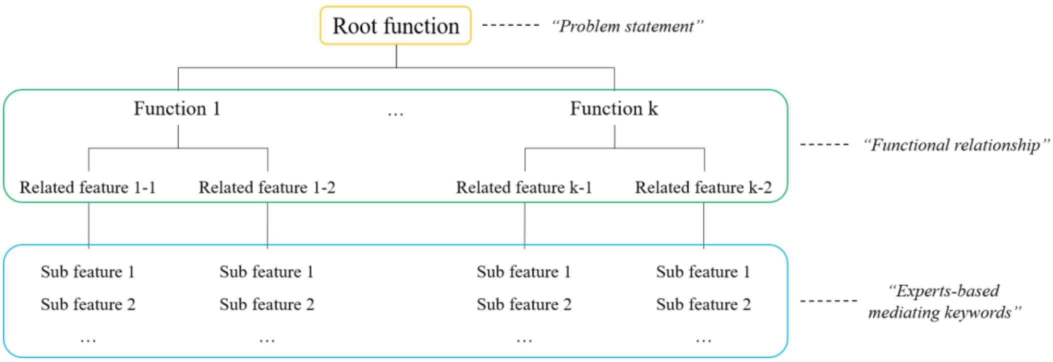

Figure 4.2 A technology tree adopted in this study

In addition, co-occurrence analysis is performed as a complementary tool to find hidden features related to the root function. In co-occurrence analysis, 'Dice coefficient' and 'Log-Likelihood' are the widely used statistical indices to judge the existence of the relationship. Based on these indices, keywords with low co-occurrence frequency are filtered and eliminated. Note that in extracting and determining data-driven keywords list, only the keywords that are not identified in the technology tree are selected. Through this, mediating keywords are constructed by hybrid use of the experts-based approach and data-driven approach. For next, all the SAO structures containing at least one of the mediating keywords are retrieved from two databases. What is important at this step is to assign which function and feature each SAO structure corresponds to, based on the technology tree, as demonstrated in Table 4.3. This task enhances the convenience of exploring and interpreting the exact solution corresponding to a specific problem.

Table 4.3 An example of retrieved results with assigned function and feature

Doc #	Problems	Related function-feature	Solutions	Related function-feature
1	(P-SAO) 1-1	FU1 – FE1	(S-SAO) 1-1	FU3 – FE2
	(P-SAO) 1-2	FU2 – FE3	(S-SAO) 1-2	FU2 – FE1
			(S-SAO) 1-3	FU4 – FE2

4.4.4 Step 4: Case Adaptation

Generally, the adaptation is to compensate for the differences between an old situation and a new one, trying to adapt an old solution to a new problem [8]. However, the ultimate goal of this study is not to find the most similar case to the problem statement. It is to identify the various issues to be considered in the problem-solving process, and to acquire the knowledge of lead users' experiences to address the potential issues proactively. It means that a guideline is required to determine critical issues from the derived P-SAOs and verify the ideas from S-SAOs proposed to resolve the issues.

Hence, in-depth analysis of the retrieved SAO structures is conducted for supporting experts' judgement. To improve a "root function", there may be a solution to various features, however, a solution to the specific feature may affect another function or related features. The more a feature that is associated with other features, the more likely it is at the center of problem-solving process. Thus, understanding these relationships can help to determine which functions and features should be solved preferentially. On the other hand, the applicability of solutions is judged by the experts based on their domain knowledge, economical efficiency, and the firm's current technical skills. The suggested guidelines are summarized in Table 4.4.

Table 4.4 Guidelines for in-depth analysis of the retrieved SAO structures

Guidelines	Description
Function level	How often does this function cause a problem? ✓ The ratio of the number of P-SAOs assigned to a particular function to the total number of P-SAOs
Feature related keywords level	Could there be a causal relationship between features? ✓ Co-occurrence frequency of the feature within the same function ✓ Co-occurrence frequency of the feature between different functions
Applicability	Checklist: ✓ Is it a solution that provides functionality meeting the concept of product? ✓ Have we tried the similar solutions? (Is it a completely new approach?) ✓ Could we gain a competitive advantage over other products through utilizing the solution? ✓ Is it reasonable in terms of product development cost? ✓ Is it accessible with our internal technological capability?

4.5 Illustrative Case Study: Drone Technology

4.5.1 Background

In order to illustrate the applicability of the proposed approach, a case study of a technology-intensive product-UAV, or so-called drone was conducted. Drone is a perfect example since it contains numerous technologies such as sensors, networks, batteries, software modules, and so on. Small changes in such technologies can have a direct impact on the performance of the drones. Moreover, drone itself holds high level of both uncertainty and opportunity as it is still in the early stage of development [69]. Therefore, not only identifying various technological issues and potential needs but also searching for a suitable solution might lead to success in the market. In this section, lead user community associated with drone is investigated, and the proposed approach is conducted in a consecutive manner.

4.5.2 Process and Results

We choose 'diydrones.com' among communities related to drone, based on the following criteria: sufficiency, diversity, and suitability. It is the leading community for personal UAVs and Robotics with over 40,000 members, where they can freely communicate their opinions. There are more than 10,000 posts that cover various topics such as aerial photography, aircraft platform, design, and software. However, it is meaningless without a text that reflects the characteristics of lead users. A typical article is shown in the following Figure 4.3. Each article contains mainly its own problem with appropriate solutions and other useful information such as title, author, and postdate. Thus, the community is suitable for performing the revised textual CBR suggested in this study.

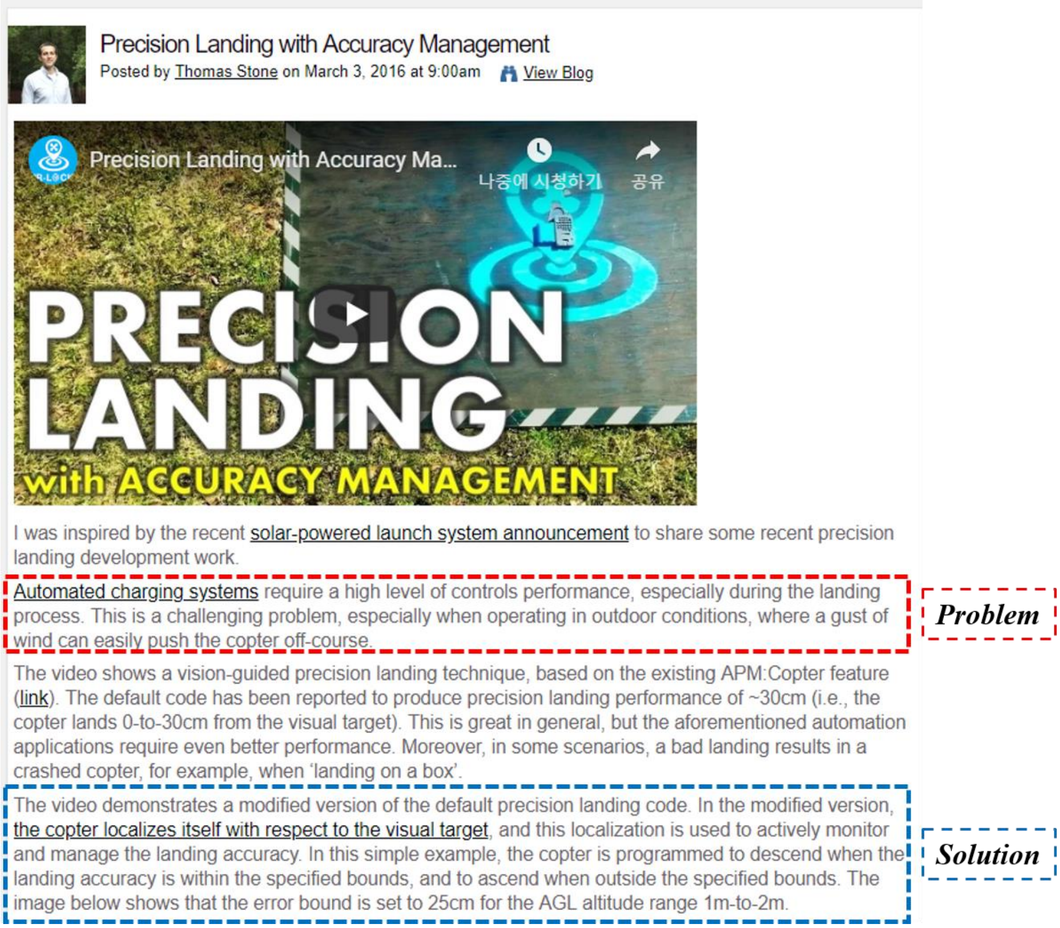

Figure 4.3 The post of lead user's experience at 'diydrones.com'

The number of articles is so large that we could not collect all of them manually. For this reason, a Java-based web-crawler was developed to download all the blog posts automatically. As a result, we collected a total of 10024 articles with the reference period from April 2011 to September 2018. Before moving to the next step, we applied Natural Language ToolKits (NLTK), which is widely used in performing the Natural Language Process (NLP) for the collected textual data. This means that pre-processing was conducted to remove the stop words such as articles, prepositions, and conjunctions. Moreover, the collected documents were parsed to separate out the parts into title, contents, author, and postdate.

Pre-processed documents are transformed to SAO structures. The total number of derived SAO structures is 385,872. According to the above-mentioned filtering process, semantic analysis and sentiment analysis are conducted to assign a certain SAO structure into a problem or a solution. Figure 4.4 illustrates an example that classifies SAO structures extracted from a certain document as problem, solution and neutral. First, the existence of synonyms is confirmed by extracting the keyword from 'Action' phrase. For example, the SAO structure, "I – was inspired by – the recent solarpowered launch system", can be regarded as a concept of the solution. Because as the first filter, the verb, "inspire" is included in the 'Action' phrase. Second, a sentiment score is calculated focusing on adverbs and adjectives for SAO structures that cannot be classified through the first filter. "a gust of wind – easily push – the copter offcourse" is the SAO structure representing a problem situation since it records a negative sentiment score of '-0.44' less than the cut off value determined by analysts. Consequentially, 17,093 P-SAO structures and 59,700 S-SAO structures, labeled with document and structure index, formed each database.

Document No.	Structure No.	Subject - Action - Object	Syntactic (1st filter)	Emotional (2nd filter)	P/S
1522	1	I - was inspired by - the recent solarpowered launchsystem	inspire		S
1522	2	Automated charging systems - require - a high level of controls performance	require		P
1522	3	This - is - a challenging problem	-	-0.2	-
1522	4	a gust of wind - easily push - the copter offcourse	-	-0.44	P
1522	5	The video - shows - a visionguided precision landing technique	-	0.12	-
1522	6	The default code - has been reported to produce - precision landing performance of 30cm	-	0.16	-
1522	7	the aforementioned automation applications - require even - better performance	require		P
1522	8	The video - demonstrates a modified version of - the default precision landing code	demonstrates		S
1522	9	the copter - localizes - itself with respect to the visual target	-	0.08	-
1522	10	this localization - manage - the landing accuracy	-	0.33	S
1522	11	the landing accuracy - is within - the specified bounds	-	0.24	-
1522	12	This approach - can be extended and customized in - a variety of ways	can		S
1522	13	the default precision landing feature - is enabled by - default	enable		S

Figure 4.4 An illustrative result of the case representation

A major issue in drone control is the difficulty of landing the vehicle on a platform [53]. Moreover, autonomous landing is one of the must-have operations of drones [10]. Autonomous landing, which guarantees simplicity, accuracy, and stability, requires a variety of requirements. However, autonomous landing is an open problem not yet definitively solved in every scenario [27]. Thus, we select landing as a "root function" of problem statement.

The mediating keywords are derived through two steps to identify the overall technological and even non-technological features associated with landing, as described in the Section 4. Firstly, a technology tree related to the root function (i.e. landing) is written by a group of experts. The group comprised five experts having experience in drone development from two years to five years. As shown in Figure 4.5, landing is implemented by functions of mobility, sensing and perception, control systems, supporting devices and others. As a result, keywords are assigned to 8 related features: power unit, driving device, aerial sensor, perception algorithm, auto control, manned control, supporting device, and etc.

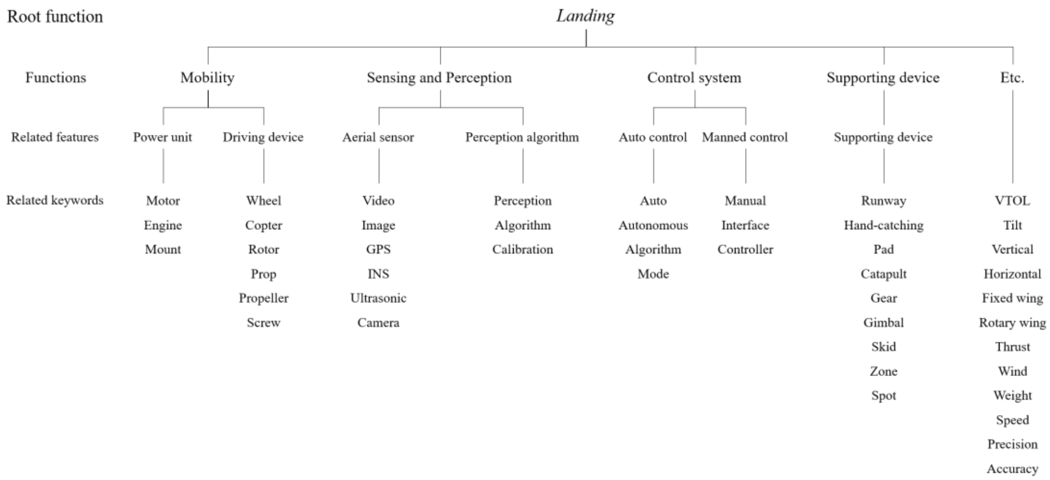

Figure 4.5 A technology tree related to the landing function

Secondly, an implementation in R with "tm" and "openNLP" packages for Dice coefficient and Log-Likelihood yields the statistical extraction of co-occurrence terms. The result is sorted so that the most significant co-occurrences are the first ranks of the list. Due to the page limit, only the top 20 terms are entered in Table 4.5.

Table 4.5 Top 20 terms extracted from co-occurrence analysis

Landing	Dice-Terms	Dice	LL-Terms	LL
1	Gear	0.2771	Gear	3151.50
2	Takeoff	0.0979	Takeoff	754.14
3	Precision	0.0601	Precision	349.86
4	Auto	0.0516	Auto	264.73
5	Autonomous	0.0419	Retractable	207.90
6	Ground	0.0381	Skids	207.02
7	Land	0.0367	Legs	164.04
8	Legs	0.0337	Ground	157.53
9	Vertical	0.0331	Parachute	152.65
10	Parachute	0.0317	Land	144.84
11	Site	0.0285	Vertical	143.47
12	Arms	0.0276	Pad	133.21
13	Altitude	0.0263	Arms	89.94
14	Camera	0.0257	Emergency	86.38
15	Motor	0.0237	Grass	85.79
16	Skids	0.0230	Height	82.21
17	Height	0.0225	Strip	76.03
18	Zone	0.0224	Frame	73.81
19	Crash	0.0217	Altitude	70.22
20	Control	0.0213	Camera	65.84

Finally, the mediating keywords list that integrates the keywords obtained from the textual data of lead users and the technology tree created by experts is shown in Table 4.6. A total of 64 keywords were determined excluding duplicate keywords.

Table 4.6 Mediating keywords for case retrieval

Mediating keywords	Experts-based keywords	Motor, Engine, Mount, Wheel, Copter, Rotor, Prop, Propeller, Screw, Video, Image, GPS, INS, Ultrasonic, Camera, Perception, Algorithm, Calibration, Auto, Autonomous, Mode, Manual, Interface, Controller, Runaway, Hand-catching, Pad, Catapult, Gear, Gimbal, Skid, Zone, Spot, VTOL, Tilt, Vertical, Horizontal, Fixed wing, Rotary wing, Thrust, Wind, Weight, Speed, Precision, Accuracy
	Data-driven keywords	Gear, Ground, Legs, Parachute, Arms, Height, Crash, Emergency, Control, Waypoint, Hover, Climb, Sequence, Throttle, Transition, Switch, Flap, Position, Frame, Site

In the next step, all SAO structures containing at least one of the mediating keywords are derived. A certain context can be inferred through a set of problems and solutions. As a result, from 1,488 documents including 'landing', 3,829 P-SAO structures and 13,122 S-SAO structures are assigned to five functions and eight features. A partial result of problems assigned to the function of sensing and perception with corresponding solutions, is described in Table 4.7.

Table 4.7 A partial result of problems and solutions assigned to sensing and perception

Reference (documents)	Requirement issue (problems)	Related features	Alternative (solutions)	Related function and feature	Innovation type
443	professional aerial photographers want to obtain detailed pictures and videos of buildings or landscapes	Aerial sensor	The dust proof coating and propulsion system protects it from external electrical components		Incremental
			The retractable landing skids make it easier to transport the drone		
1241	serious real time video processing power is required on the order of significant multi GPU based system	Aerial sensor	The Exo360 brings together the latest in drone and VR technology		Disruptive
			The drone can capture video in both linear and spherical modes	Sensing and Perception – Aerial sensor	

1682	Most camera will not operate for more than 2hours without external power supply	Aerial sensor	The major breakthrough is the new Lithium Ion battery pack rated at 3C continuous discharge		Disruptive
	The camera will need external power for the long flight duration		the new long range conservation drone open up the possibility to map or video area	Sensing and Perception – Aerial sensor	
3045	The crucial component for geotagging images is the GPS receiver	Aerial sensor	images georeferenced using common GPS modules are sufficient image matching techniques might be applied the camera settings have to be adjusted for every single image		Incremental

	the problem is a proper calibration of the gimbal	Perception algorithm	NADIR gimbals provided an easy way to reduce perspective distortions	Supporting device – Supporting device	
			The offset can easily be determined using Google Earth		
3863	my photos came out blurry	Aerial sensor	The landing gear also allows you to take off like a traditional plane	Supporting device – Supporting device	Incremental
			Canon Hack Development Kit allow you to use intervalometer scripts to trigger camera	Sensing and Perception – Perception algorithm	
3963	you want to see a really impressive use of image processing		It used the simulated copters position and altitude	Etc	Incremental
			It created a simulated image of the horizon and the balloon	Sensing and Perception - Aerial sensor	

3971	you want to make videos that look drop dead awesome	Aerial sensor	The AirDog is designed to withstand impact		Disruptive
			The prop guard, an optional accessory provides further protection	Mobility – Driving device	
	we need much higher level sensors and very complicated sensor fusion algorithms	Perception algorithm	The camera is recording with flawless accuracy		
			AirDog uses axis gyro-stabilized gimbal	Supporting device – Supporting device	
5385	the best possible airspeed measurement needed to calibrate the sensor	Perception algorithm	The first change is to add support for a new digital airspeed sensor	Sensing and Perception - Aerial sensor	Incremental
			a new digital airspeed sensor is a huge advance over the analog sensors		
			a sensor with low thermal drift is very nice		

			a small state Kalman filter can automatically calibrate the airspeed ratio	Sensing and Perception – Perception algorithm	
			The idea is to make APM tuning parameters for an airframe be the same at sea level		
5866	We need to manually transfer the GPS waypoint	Aerial sensor	We triangulate the gunshot source and provide a GPS coordinate	Sensing and Perception - Aerial sensor	Incremental
			it generates a UAV script to APM planner		
5896	The first version of MatrixPilot suffered from GPS latency	Aerial sensor	This precision navigation is the result of several recent innovations in the MatrixPilot algorithms	Sensing and Perception – Perception algorithm	Incremental
			The main innovation is using IMU course over ground	Etc	

			We totally eliminated the use of angles	Etc	
			the navigation computations is now based on vectors or matrices	Sensing and Perception – Perception algorithm	
7143	stabilized camera gimbals suffer from both lag	Aerial sensor	position control actually does make more stable flight possible	Etc	Disruptive
			the horizontal motors will completely take over horizontal position	Mobility – Power unit	
			the Vertical thrust motors will be responsible for stabilization altitude control		
			effectively the air frame yaw could be used for camera pointing	Sensing and Perception – Aerial sensor	

			it should greatly reduce the further requirements for camera stabilization		
8060	the camera has a small field of view	Aerial sensor	A horizontal firing camera of higher resolution and downward firing sonar for altitude might do it	Sensing and Perception – Aerial sensor	Disruptive
	The camera would have to detect optical flow of vertical objects				
	modern object detecting algorithms aircraft only use optical flow	Perception algorithm			
	the camera needs a complete view of the axis of rotation	Aerial sensor	A fisheye lens pointed straight down seems the only way	Sensing and Perception – Aerial sensor	
			A fisheye lens would solve everything		

			the idea of sonar for altitude & a horizontal camera for position continues to haunt		
9548	the camera will need to handle very heavy vibrations	Aerial sensor	The real time tracking system is used in this technology		Incremental
			this technology is developed by Vicon Motion Systems		
			the right visibility markers are cheap enough for a hobbyist		
	the camera needs to be somewhat durable as soft landings	Aerial sensor	we have developed a cost effective camera	Sensing and Perception – Aerial sensor	
			our higherend cameras is available at a more affordable price	Sensing and Perception – Aerial sensor	

Having retrieved problems and solutions with assigned function and feature, we then performed an in-depth analysis. The approximate importance is calculated at the function and feature level according to the guidelines suggested in Section 4 as follows. The number of P-SAO structures assigned to the five functions is shown in Table 4.8. Landing is basically conducted based on landing point acquisition and obstacle detection. Thus, it can be interpreted that importance of the function, 'sensing and perception' in software aspect and importance of the function 'supporting device' in hardware aspect are emphasized.

Table 4.8 In-depth analysis of 'Landing' (function level)

Function level	Mobility	Sensing and Perception	Control system	Supporting device	Etc.
Number of assigned P-SAO structures	719	1147	270	932	761

In more detail, Table 4.9 shows the results of in-depth analysis of feature related keywords to determine which features have a crucial role in the 'sensing and perception' function.

Table 4.9 In-depth analysis of 'Sensing and Perception' (feature level)

Sensing and Perception (function)	Features and related keywords		
Co-occurrence frequency within same function	Aerial sensor	Video	152
		Image	293
		GPS	92
		INS	7
		Ultrasonic	11
		Camera	507
	Perception algorithm	Perception	42
		Algorithm	156
		Calibration	78
Co-occurrence frequency between different functions	Aerial sensor	Video	12
		Image	43
		GPS	172
		INS	3
		Ultrasonic	7
		Camera	266
	Perception algorithm	Perception	23
		Algorithm	177
		Calibration	31

4.6 Discussion

4.6.1 Incorporation of Lead User Knowledge

This paper proposed a novel approach to the textual CBR to identify problems and explore feasible solutions from knowledge in lead user communities, addressing a lack of approaches to leverage the big textual data therein. It embraces innovation from outside the boundaries of the firm, especially lead user innovations, into their own innovation strategies in two ways: incremental innovation and disruptive innovation. First, incremental innovation can be achieved by improving the performance of existing products based on the lead users' uncomfortable experiences. Second, totally new needs can be utilized as a starting point for disruptive innovation.

In particular, the case representation and case retrieval process, which have been emphasized in textual case-based reasoning, are further improved in this study. In the process of conducting our case study, the effectiveness of the proposed approach is confirmed, as shown in Table 4.10. First, the classification of SAO structures in a document into problem, solution, and neutral statement facilitates the interpretation of case retrieval results. The context can be grasped through SAO structures beyond the keyword based approaches. Second, the experts' domain knowledge and the lead users' knowledge based on their experience are used complementarily in the form of mediating keywords. It extends the scope of retrieval process exploring various solutions related to the problematic root function.

Table 4.10 The effectiveness of the proposed approach compared with the keyword-based approach

Illustrative document: Drones have already been in use in the security and surveillance industry, bringing a significant change in how the operations are carried out. However, most current aerial security and surveillance systems are either tied to a particular drone hardware, or need significant manual intervention during operation. These solutions lack critical features and software capabilities, such as, AI and machine learning for automated alerts, automatic mission scheduling, compatibility with wide-range of drone hardware, etc. This makes it expensive, and often infeasible, to deploy the drone-based security/surveillance solutions at scale. Drones have already established the value that they bring to the table, in terms of mobility, unrestricted bird's eye view and accessibility. The focus is now on efficiencies and realising a meaningful return on investment for wide commercial adoption. This calls for integration of "intelligence" and "connectivity" with drones, to build completely automated and integrated workflows. FlytSecurity offers a plug-and-play, drone-agnostic, SaaS platform to quickly deploy and scale drone-based automated security operations. This significantly cuts down the cost of development and time to market, translating into an attractive ROI for the drone security service providers. With a wide range of features, like, 4G/LTE connectivity over unlimited range, live video, control and telemetry, fleet management (for simultaneous coverage of a large, distributed facilities), AI/ML for automated alerts, automated mission schedules, FlytSecurity enables fully-automated 24×7 operations at scale. Compatibility with any drone hardware, further makes FlytSecurity easy to adapt to variety of customer requirements (large/small drone, long/short endurance, quad-planes/multicopters, thermal/RGB sensor, etc.), and makes it easy to upgrade hardware at any time.

SAO-based approach	Keyword-based approach
Problem/Solution/Neutral SAO structures - most current aerial security and surveillance system are either tied to a particular drone hardware (N) - a particular drone hardware need significant manual intervention (P) - These solutions lack critical features and software capabilities (P) - This makes it expensive and infeasible (P) - The focus is now on efficiencies (N) - This calls for integration of intelligence and connectivity (P) - FLytSecurity scale drone-based automated security operations (N) - This significantly cuts down the cost of development and time (S)	DocumentTermMatrix (TermFrequency) - drone (9), automate (5), security (5), operation (4), hardware (4), scale (3), surveillance (3), access (2), alert (2), compatibility (2), connectivity (2), deploy (2), integration (2), mission (2), range (2), solution (2), mobility (1), plug (1), telemetry (1) DocumentTermMatrix (Tf-idf) - FlytSecurity (0.3572), automate (0.1295), security (0.1156), surveillance (0.0894), alert (0.07767), AIML (0.0715), connectivity (0.0661), operation (0.0639)

In order to verify the effectiveness, the comparison of the retrieved results is conducted by performing the traditional keyword-based textual CBR on the same database. The problem statement, 'autonomous landing accurately and reliably on a moving platform', is converted to a keyword vector as an input of textual CBR. The input keywords are autonomous, landing, takeoff, accurate, precise, reliable, moving, platform, ground, and station. Then, the collected documents are transformed to term-document matrix for calculating the cosine similarity. As a result, the keywords list, which represents the documents in the order where the similarity score is high, is retrieved, as shown in Table 4.11. It is inevitable to review the document to confirm the exact solution since the keywords list alone could not deliver even the concept of a solution. In addition, it is nearly impossible to explore issues comprised of keywords other than the input keywords.

Table 4.11 A keywords list of the top 10 documents in the order where the similarity score is high

Document no	Similarity score	Keywords
9283	0.2108	landing, speed, camera, height, helicopter, object, platform, surface, vision, chip, hover
5406	0.2107	autonomous, landing, compass, gps, multipath, back, ground, throttle, wind, runway
4904	0.2107	landing, car, roof, quad, antenna, moving, gear, airbone, speed, control, technology
4822	0.2107	landing, flap, fuselage, canopy, gear, skid, gimbal, grass, wing
7475	0.2013	landing, gear, camera, polycarbonate, mount, platform, camera, smart, simulator
2546	0.2009	autonomous, landing, plugin, platform, infrastructure, LTE, solution, environment, scale

1407	0.2009	landing, code, copter, rangefinder, altitude, precision, accuracy, platform
1164	0.2009	autonomous, video, Bluetooth, autopilot, data, online, accurate, cloud, device
395	0.2003	landing, wing, Bluetooth, autopilot, , wind, perching, algorithm, bird, disaster
1828	0.1924	autonomous, landing, data, time, altitude, thrust, algorithm, solution, mapping, sensor

4.6.2 Combination of Other Methods

This study contributes to the field in that it presents a new data source for supporting development of technology-intensive product. As shown in the case study, there are heterogeneous problems and adequate conceptual solutions in lead user communities. The use of CBR is clearly a methodology that reflects the distinctive characteristics of lead user communities. However, it is also important to select the prerequisite issues where the technology-intensive product is in the early stage of growth. This is closely related to the case adaptation step of the CBR. For this reason, it seems to be required to combine a certain method with other powerful methods. For example, it should be noted that other methods in Table 4.12 can be employed as remedy to develop quantitative indicators for the case adaptation guideline suggested in this paper.

Table 4.12 Possible methods applicable to the case adaptation stage

Method	Applicability for development of case adaptation guideline
Topic modeling	Topic modeling technique such as Latent Dirichlet Allocation (LDA) defines topics based on the distribution of keywords from a large corpus and classifies documents into each topic. Specifically, it is possible to calculate the importance of features through the value (i.e. beta in LDA) of the influence of the keyword on the topic. Thus, it helps to judge the technical features that are central to a particular problem.
Network analysis	A network can be constructed based on the co-occurrence matrix between keywords. In order to observe the changes of a certain keyword in the network, degree centrality, closeness centrality, and betweenness centrality indicating the structure of the network can be utilized. They can serve as indicators for evaluating importance of technical features.
MCDA techniques	The process of selecting the problem to be solved should take into consideration several criteria rather than one. The multi-criteria decision analysis techniques fit this situation. Typically, AHP and ANP can reflect the firm's strategic objectives in the case adaptation process.

4.7 Conclusion

This paper proposes a revised textual case based reasoning approach to reflect lead user's characteristics in technology-intensive product development. The utilization of lead user community is improved by separating the problem and solution SAO structures from a post shared by the lead user in case representation stage. Where the problem statement is unclear in case retrieval stage, we suggest a hybrid approach of technology tree and co-occurrence analysis, whose main focus is to identify technological structure of the product and potential problems. It enables us to explore various problems and exact solutions with expert judgement at last.

This paper contributes to research in three ways. First, from the theoretical perspective, we propose a differentiated approach reflecting the characteristics of the lead user. The elaborated textual CBR extends the application area of lead user communities into an incremental or even disruptive innovation of product. Second, from the methodological perspective, this paper focuses on SAO structures instead of keywords to minimize the loss of information and employs technology tree as a remedy for providing a broad overview of issues in technology-intensive product development. Third, from the managerial perspective, this study provides an organized approach for embracing lead user innovation from outside the boundaries of the firm into their own innovation strategies. Since our approach can both efficiently identify the problems from lead users' experiences and also suggest ways to find out related solutions based on their ideas, firms can benefit by incorporating it to their R&D efforts.

Despite this contribution, there are limitations to be addressed in future studies. First, in the case adaptation stage, we provide only rough guidelines to find critical functions and features to be addressed. The applicability of solutions is evaluated through the suggested checklist. These processes could be further improved by using several quantitative approaches. So another fruitful avenue for future research would be to employ relevant methodologies to automate the evaluation process. Second, the collected documents contain additional information such as author, postdate and user comments. Thus, other research topics can be performed utilizing the additional information in two ways: time-series analysis and impactful lead user identification. There have been a lot of works to determine critical and peripheral functions of a product from customers using a market analysis approach such as Kano model. Identifying the key needs through real-time monitoring is more appropriate for the first stage of demand-pull innovation than the static analysis in rapidly changing markets. Moreover, interaction among lead users occurs in the form of discussion and user comment. Utilizing these interactions in the network analysis with quantitative indicators such as the number of views and comments on the post, impactful lead users can be identified. This work can help firms select target customers for qualitative approaches such as focused interview and survey. Third, it utilized the single source of lead user community as an experimental study. However, the increasing complexity of technology-intensive products is highlighting the role of collective intelligence. There has been an attempt to absorb various knowledge sources in academia and practice. In order to utilize more various sources of lead user communities, it is necessary to develop a process such as group fuzzy cognitive maps (G-FCMs) to merge into integrated knowledge.

Chapter 5

Planning for Data Innovation

5.1 Introduction

It is widely anticipated that there has been and will be drastic paradigm shift towards the data-driven world. Coupled with the advances in analytic technologies, data becomes critical in the business systems. People, machines, and systems are connected via online, communicating based on the thousands of data generated by IT infrastructures. The notion industry 4.0 has changed industry landscape towards predictive and proactive analysis using real-time data processing.

This is especially true in smart service development where value is created and extended based on huge amounts of data. Smart services require significant efforts to integrate devices, services, and technologies. What is critical in integrating all these elements is data. Data works as an important player in the smart service ecosystem. Data enables smart service components be connected.

Therefore, data should be considered as an important part of strategic technology management. In other words, what kinds of data should be generated and how they should work in the system is a critical question to the smart service development. In addition, what kinds of data should we have is another important question. This means data should be considered not only as a means for strategic planning, but also as an important subject to be planned. In response, a systematic framework and methodology to integrate data into the strategic technology management is required.

For this purpose, a technology roadmap worth the effort. The technology roadmap, since its invention in 1980s, has been a prominent technology planning tool at the hands of strategic planners and researchers. It has been widely used in practice due to its power for linking markets, products, and technologies over time [44]. Since its popularity in supporting both technology push and market pull in strategic planning, many previous works have been conducted to use technology roadmaps in a variety of applications [43, 56]. Cannata et al. [18] tried to develop a technology roadmap for supporting intelligence system in manufacturing. Lee et al. [76] developed an integrated roadmapping process for smart city development, which integrates services, devices, and technologies using quality function deployment.

While a number of studies have been conducted to widen and deepen the application area of technology roadmaps, the research direction should be turned to highlight the role of data and properly integrate data into the strategic planning process. Confronted with the importance of big data in strategic planning, many researchers tried to integrate data into the technology roadmaps [43, 56]. Geum et al. [43] focused on how data can be used in technology roadmapping, and suggested a data-driven technology roadmap using association rule mining. Jeong and Yoon [56] used patent information to the technology roadmaps by analyzing structural and temporal patterns of patent development. Yu and Zhang [131] developed a patent roadmap to support proper understanding of current patent competitive situation and perform relevant patent layout planning. Jin et al. [57] also used patent data to the technology roadmapping by applying text mining and quality function deployment.

However, previous studies simply employed patent database or product/service database to analyse the trend of technologies or market. Massive patent/product/service data is collected, analysed, and used to illustrate how technological and market trends go on, which can be further used for evidences for strategic planning. This means previous studies mainly deals with roadmapping 'by' data. They view data as ex post, which means data is not the subject of planning. Rather, data can be effectively used to the strategic planning afterwards.

At the same time, when companies try to develop new products or services, data should be considered as the subject of planning, which views data as ex ante. This means technology roadmap should be conducted 'for' data, not 'by' data. However, literatures to view data as ex ante are surprisingly sparse. It still remains a void in the literature to deal with how data works in strategic technology planning, and how data should be properly integrated in the roadmapping process. Most importantly, current literatures view data simply as an intermediate for strategic management, not viewing data itself as a subject to be systematically planned.

In summary, data should be one of the subject of strategic planning, which means data layer is required in the technology roadmapping process. Taken together, this study suggests a concept of data-integrated technology roadmap, and proposes relevant structure, typology, and roadmapping process to show how to integrate data into the technology roadmapping. To accommodate the different situations that data works in the smart services, this study firstly suggests the types of data integration based on literature review and practical business cases. Second, concept and structure of data-integrated technology roadmap is suggested, adding a data layer as intermediate and functional link for planning corresponding smart services. Third, typology of data-integrated technology roadmap is also suggested considering types of data integration, in order to provide flexible use cases in different situations.

The remainder of this research is organized as follows. Literature review deals with how data is used in recent roadmapping studies, as well as the composition of smart service system focusing on data. Next, types of data integration are suggested with typological approach. Following on the data integration, the concept and structure of data-integrated roadmap is suggested. The typology of data-integrated roadmap is also provided with their different characteristics. To illustrate how our data-integrated roadmap works, a simple case example is provided. Finally, the contribution of this study along with future research directions is also discussed.

5.2 Roadmapping Process for Integrating Data

Since data becomes central in the business environment, substantial number of studies have been conducted to employ data to the technology roadmapping process. Therefore, techniques and methods to support data-driven technology roadmaps have been extensively suggested [43, 56, 57, 131].

Since technology roadmaps generally aim to plan future technologies or products, patent information has been widely accepted for data-driven roadmaps. Jeong and Yoon [56] have employed patent information to the development of technology roadmaps by analyzing structural and temporal patterns of patens. Jin et al. [57] also used patent data to the technology roadmapping by applying text mining. Zhang et al. tried to conduct SAO analysis for patent documents, and use this information to find problem & solution patterns in order to support technology roadmaps. Geum et al. [43] collected the explanation of mobile application data from Appstore, and used this for technology roadmapping.

While previous studies have valuable implications in integrating and utilizing data into the roadmapping process, previous works still remains 'roadmappig by data'. This means the result of data analysis is used as the guideline for decision making of product or service planning, by providing objective evidences for product trends or service trends. Even if these tasks worth the efforts, what is required in practice is to consider the role of data in the strategic planning process, and consider data as the subject of planning. In response, 'roadmapping for data' should be conducted, which means new roadmapping process is required to consider the data as an important asset as shown in Table 5.1.

Table 5.1: Distinctive perspective of this study

	How roadmaps work	Characteristics	Concepts
Previous work	Roadmapping "*by*" data	Tactic	Ex-post
Our study	Roadmapping "*for*" data	Strategic	Ex-ante

5.3 The Role of Data in Smart Service Systems

Prior to develop data-integrated technology roadmaps, understanding how data works in smart service systems is very important to make a structure and typology for smart service roadmaps.

5.3.1 Macro level

Considering generation (data source: where does the data come from?) and flow of data (data interaction: where does the data affect?), smart service system mainly consists of three key elements, device, user, and smart environment as shown in Fig 5.1. In this context, a device is utilized as a delivery platform of services, users can be both a service consumer and even data provider, and the smart environments including sensors, micro-processors, network technologies, etc., connects devices and users with vast amount of data.

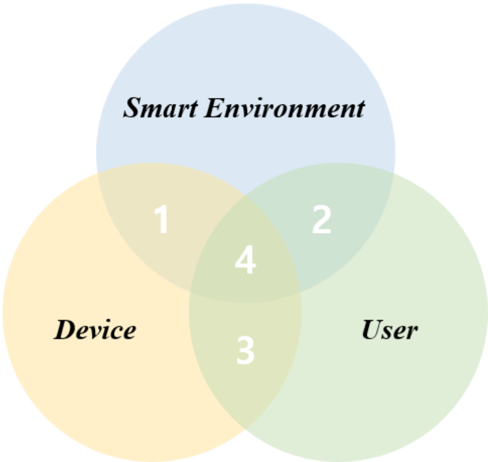

Figure 5.1: Structural aspect of smart service system (Macro: system level)

According to the data source and data interaction, the four intersection cases can be described as the following Table 5.2. For example, in a real-time bus route information system, data interaction between smart environment and bus enables the system predict the arrival time at the bus station (intersection area '1'). The information is delivered to the public through the panel at each bus station. As another example of navigation systems, optimal route was explored based on the embedded static map data in the device without regard to real-time traffic congestion in the past (intersection area '3'). On the other hand, due to the smart environment and the development of data technologies, recent navigation system shows the estimated arrival time along with optimal route based on real-time traffic congestion and prediction algorithm (intersection area '4'). In other words, the smart service system is realized through generating, collecting, and analyzing different types of data and interacting among the elements.

Table 5.2: Four intersection areas according to data source and interaction

Intersection	Data source	Data interaction
1	external data generated from device	smart environment -> device smart environment <-> device
2	external data generated from user	smart environment -> user user -> smart environment
3	generated from device generated from user	device -> user user -> device
4	external data generated from device generated from user	device <-> user <-> smart environment

5.3.2 Micro level

While the structural aspect of smart service systems with data source and data interaction is addressed in the macro level, the micro level deals with the technological aspect of services focusing on the data. Ultimately, the value of a smart service system depends on how the system handle the big data. Thus, hardware and software technologies required for delivering new values are the key components of the service as shown in Figure 5.2. As emphasized, big data is a crucial element for 'smartization' of services and often it serves as a starting point of creating new values. In this situation, hardware technologies including electrical and mechanical parts of a device mainly focus on detecting and collecting data. Meanwhile, software technologies such as embedded operating system, algorithms and cloud storage are needed to store, manage, and analyze the big data in and out of the devices.

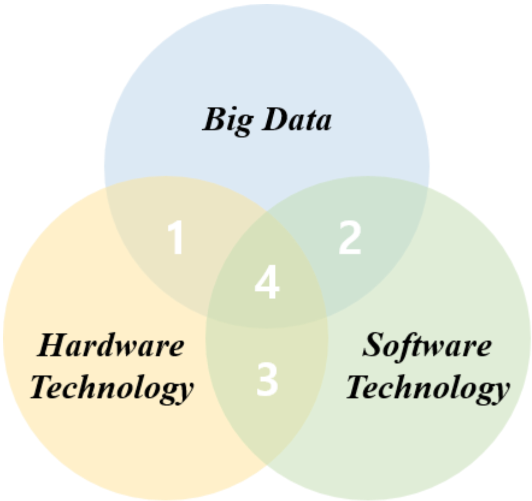

Figure 5.2: Technological aspect of smart service system (Micro: service level)

According to the purpose of technologies based on big data perspective, the four intersection areas can be interpreted as the following Table 5.3. Previously, research on PSS have highlighted the hardware and software technology itself. In the case of smart service systems, however, the importance of technologies related to 'how to collect data efficiently' and 'how to utilize data effectively' is emphasized. As a result, the acquisition of big data with an application of data technologies creates and delivers distinctive values for customers.

Table 5.3: Four intersection areas according to the purpose of technologies

Intersection	Interpretation
1	Data collection and storage - Detection and collection of the naturally generated data requires new hardware technology development: market-pull & data-pull - Accessing to new types of data that can be collected through the developed technologies: technology-push
2	Data storage, management, and analysis - Requiring new software technology development (e.g. image and voice recognition algorithm) to handle entirely new and unprecedented types of data: market-pull & data-pull - Exploring data that the developed algorithms can be applied: technology-push
3	Data transmission, connection, and synchronization - Increasing service consistency in various context by synchronizing information with the devices
4	New value and enhanced value - Improvement of existing service system (incremental innovation) or generation of completely new smart service system (disruptive innovation) through big data with technologies

Note that our typology in data role is developed based on the consideration that how data influence the technology roadmapping process, investigated in both macro and micro level above. How data is involved in smart service systems can be characterized in Figure 5.3: data as supporter, data as mediator, and data as value generator.

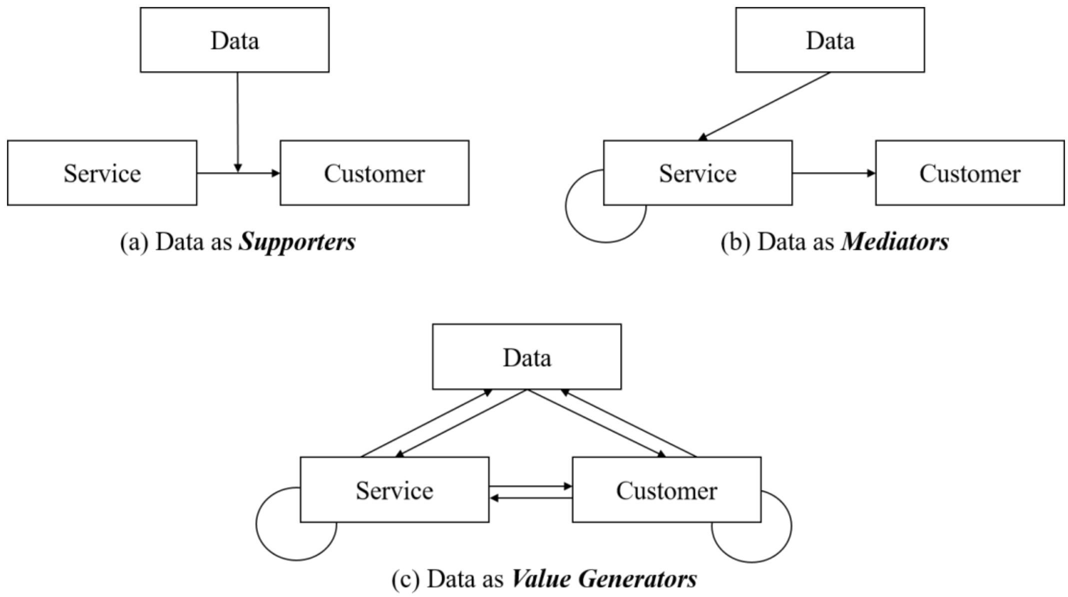

Figure 5.3: The role of data in smart service systems

First, when data simply supports and facilitates the existing service process, it is called 'data as supporter.' In this type, data supports the communication and processes between service and customers. For example, when blood testing results are delivered not via the hardcopy papers, but via mobile application that enables to monitor the increase and decrease of specific indicators, it can be called data as supporter. This service, providing the results of blood testing, can be provided without the data driven mobile services, which denotes the role of data in this paper is simply supports and facilitate the service process.

The second type is data as mediator, which means data works as the means of service optimization. Which means data itself is used for enhancing, improving, and optimizing service processes and outcomes. Therefore, service becomes intelligent in terms of service process management. For example, when blood pressure and pulse is collected and analyzed over time, and outlier detection is performed to check whether irregular or unexpected patterns are found, this signal is automatically transferred to the hospital or nursing call is performed. Since data mediates the service process to be optimized, we can call this as 'data as mediator.'

The last type is called 'data as value generator,' which is the most evolved in terms of utilizing and managing data-related communications. In this type, data is working as customer optimization, which has evolved from service optimization. When data is working as value generator, new value is provided to customers by satisfying customers' own needs and environment. When data is generated from patients from their skin or body, this is analyzed for customer optimization purposes. Then, current services are improved or enhanced based on customer optimization. For example, temperatures or humidity is adjusted according to the patient health condition, or preferred prescription or diagnosis is automatically provided by optimizing customers' condition. Since customers can continuously provide their needs and data from the smart service system, it is two-way communication services in terms of data-customers, data-services, and services-customers. Table 5.4 summarizes how each type differs according to its data, service, and customers.

Table 5.4: Data integration type according to the technology intelligence level

	Data	Service	Customer	Intelligence level
Data as supporter	Working as process input	One-way communication	Getting services using data	Low
Data as mediator	Working as service optimization	One-way communication	Getting services and providing selective feedback to services	Medium
Data as value generator	Working as customer optimization	Two-way communication	Getting services and providing autonomous feedback to services and data	High

5.4 Data-integrated Technology Roadmap

5.4.1 Overall Structure

The generic structure of suggested data-integrated technology roadmap is explained in Figure 5.2. Compared to the generic technology roadmaps, we added one new layer – data layer – into the data-integrated roadmap. This is because data should be the 'subject' of strategic planning in the data-driven worlds, not simply working as providing ex-post evidences.

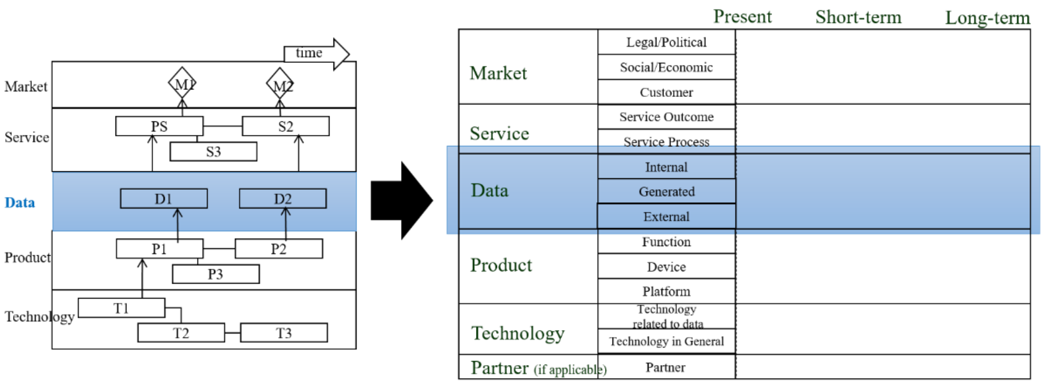

Figure 5.4: The generic structure of data-integrated technology roadmap

The left side of Figure 5.4 illustrate how data layer is inserted in the traditional technology roadmap. Note that the position of data layer can be changed according to the roadmapping processes. What is core in this figure is that data should be the target to be planned in the strategic planning process. The right side of Figure 5.4. shows how each layer of technology roadmaps can be developed for each specific layer. Note that each layer is composed of several sub-layers. As other layers, data layer is also composed of three sub-layers: internal data, generated data, and external data. The internal data is defined as data that is owned and prepared by company itself. The generated data is defined as data that is generated from service processes or customer behavior. Finally, external data is the data that is obtained from outside the company. The reason that we have to consider three layers simultaneously is to illustrate how different data sources can be collaborated and communicated in a smart service system, and to indicate how functions, devices, platforms, and services are affected from different types of data. In addition, technology layer is again divided into two major parts: technology related to data, and technology in general. This is because the term ‘technology’ in data-integrated roadmap seems ambiguous. In some cases, data cannot be properly analyzed or utilized just because there are no relevant technologies to be employed. To reflect that technology related with data analytics is sometimes critically required to the new products/service planning, we added this layer. Note that partner layer is also added. When external data is required, this means firms have to make alliances or collaboration to get desired data from outside the company, which means planning of partners is also the subject of strategic planning.

5.4.2 Typology and Roadmapping Process

Even if generic structure of data-integrated roadmap exists, it requires different roadmapping process depending on the role of data in a smart service system as shown in Table 5.5. In the case of ‘data as supporter’ and ‘data as mediator’, roadmapping starts with the technology layer for technology-push innovation or the market layer for market-pull innovation. The important thing to be considered in the process is about ‘data’. Where the data serves as a supporter, the overall process is the same as creating an existing PSS technology roadmap. At final step, which data to provide additionally in the system is determined, as shown in Figure 5.5 (a).

When data acts as a mediator, the first step of roadmapping process is to explore the changes and trends in market requirements or to consider the development and trends in technology. A typical example is the need to improve the intelligence level of existing services due to the rapidly changing market trends. To address this, the existing data should be linked with required new types of data for continuous and sustainable innovation of that service. In this process, technological requirements are coupled with data and product/service layer, as shown in Figure 5.5 (b). Lastly, if data works as a value generator, the first thing to do is what value can be created and delivered utilizing the data. After that, it is linked to product/service and technology layer, as shown in Figure 5.5 (c). Let's imagine that there is data on the usage location and time of asthmatic patients through the nebulizer's GPS sensor. This unprecedented data can serve as a starting point of new business such as an improvement of navigation system helping asthmatic patients not to pass through dangerous areas.

Table 5.5: Typology of roadmapping process according to the role of data

Role	Roadmapping starts with	Data layer roadmapping
Data as Supporter	- Technology layer (for technology-push) - Market layer (for market-pull)	Last
Data as Mediator	- Technology layer (for technology-push) - Market layer (for market-pull)	Middle
Data as Value generator	- Data layer (for data-push)	First

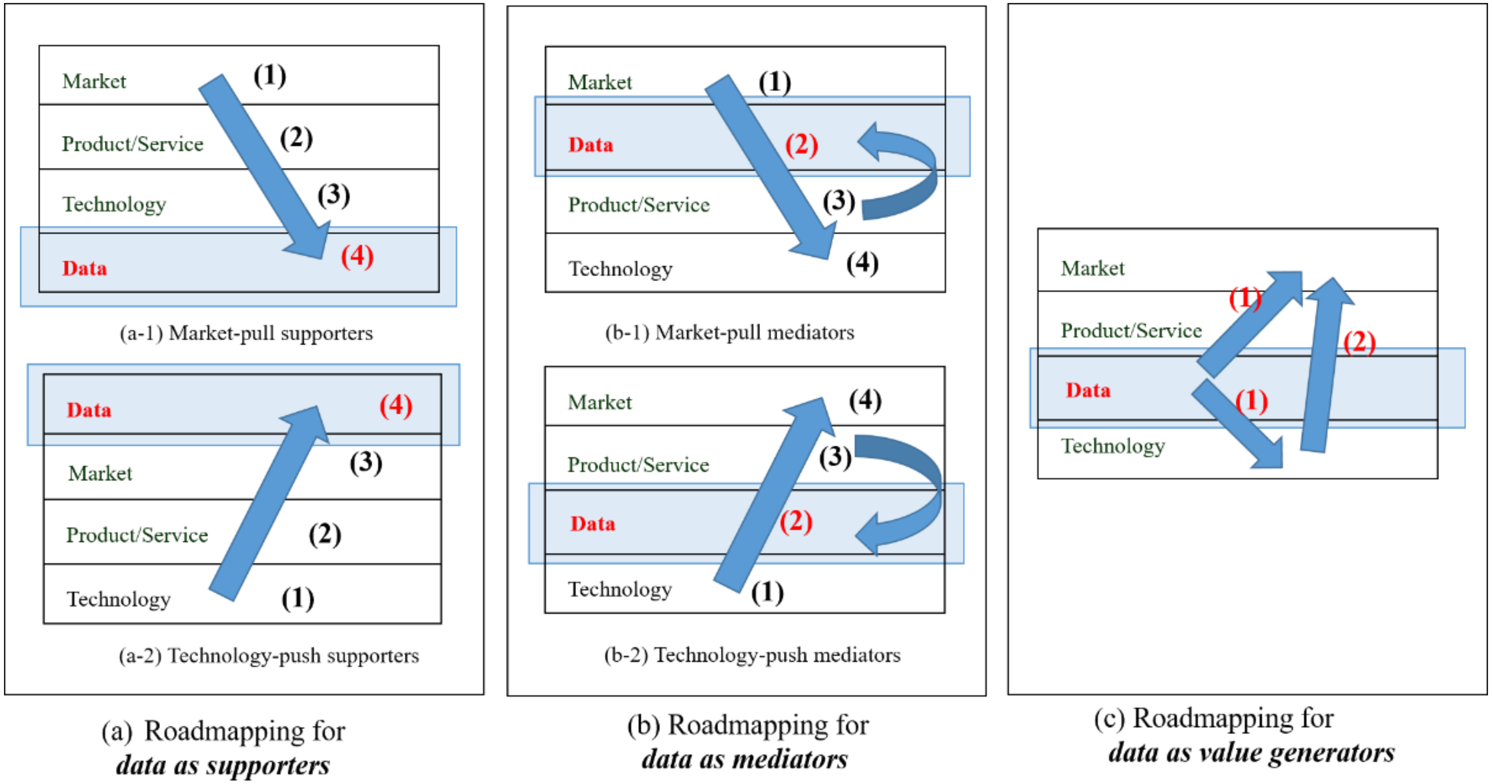

Figure 5.5: Roadmapping process of each type

5.5 Discussions and Conclusions

5.5.1 Establishment of an Analytical Perspective on Smart Service System

Based on the data integration types described in the Table 5.4, cases of the education service system can be interpreted as follows. First, educational video contents enabled the distant learning services in the education industry, where contents were provided only via person to person in offline spaces. In this context, the educational video contents work as supporter. Second, the service that monitors learning process of the user based on the user's learning log data, and establishes the education schedule, as well as provides the notification of learning plan is the case of 'data as mediator'. Third, there is a service that analyzes problem types and related contents which users have difficulties in solving, and then automatically provides the appropriate set of problems for iterative learning. That is, data is utilized as a value generator in the process of providing user customized education service. Characteristics of the exemplified education service systems according to each data integration type can be summarized in the following Table 5.6.

These data integration types can serve as guidelines for analyzing a certain smart service system. Specifically, several insights can be provided to the following questions at the planning stage of the smart service system: (1) what kinds and types of data is appropriate and required, (2) what is the role of users as a data provider in the system, (3) what to interact with through data. This process help practitioners increase understanding of data sources and data interaction in macro level (system level).

Table 5.6 Characteristics of the education service systems

	Data	Service	Customer
Data as supporter	Education video contents	Education service is delivered to users via online	Getting education contents remotely
Data as mediator	User's learning log data	Service functionality is diversified and the learning schedule (i.e. service delivery plan) is optimized	Getting not only education contents remotely but also mobile app notification service is provided
Data as value generator	Historical data of user's problem solving cases	A user-customized and optimized education contents are automatically identified	Getting education contents remotely and providing historical data to the service system continuously and automatically

5.5.2 Visualization of Planning for Smart Service System Deployment

The proposed structure of data-integrated technology roadmap visualizes the role of data for planning a certain smart service system by adding a data layer as intermediate and functional link at micro level. The inclusion of data requirements in the ‘data’ layer help us clarify the process of planning related hardware and software technologies as shown in Figure 5.6. The autonomous feeding service can be planned by linking with new kinds of data and related technologies based on video data.

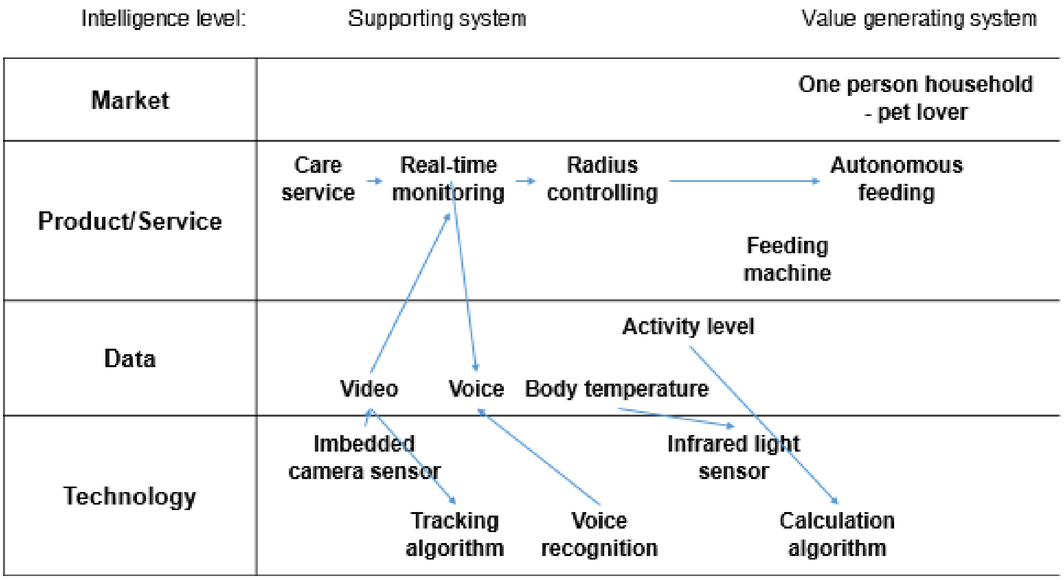

Figure 5.6 An illustrative example of data-integrated technology roadmap

It reflects not only technology-push and market-pull perspective, but also includes the direction of data-pull or data-push in terms of a new innovation driver. We define 'data-pull' as the process of identifying what kind of data should be utilized to obtain exact information requirement to provide a particular intelligence level of service. On the other hand, 'data-push' is the process of determining which hardware and software technologies should be applied based on the currently available data. Consequently, both data flow and service flow are considered in planning new smart service systems in the data-integrated technology roadmap. It supports service managers to communicate efficiently with colleagues by visualizing the direction of the smart service system development.

Chapter 6

Concluding Remarks

This thesis has endeavored to organize the development and roadmapping process of smart service systems considering their distinctive natures – a systematic process of service innovation planning based on quantitative and qualitative approach. We have greatly expanded the knowledge boundary horizon by incorporating online user communities, such as mobile app service review data and lead user community, and have further developed theoretical and methodological frameworks for the proactive planning process of novel smart service systems. The process comprises planning for system, technology, and data innovation.

As mentioned earlier, this research is aimed at addressing the following questions: (1) Which components to be considered and how the components are determined when designing a smart service system? (2) What kinds of data sources are useful and which techniques are suitable for improving existing smart service systems or exploring new smart service systems? (3) How could we plan and communicate new smart service system concepts in a visualized form? The primary contribution of this thesis lies in systematic attempt to gain insights into the utilization of online user communities (e.g. lead-user communities and review websites) and text analytics in the field of service innovation planning. In particular, each study has made the following contributions for theory and practice, as listed in Table 6.1. The detailed implications of each chapter are summarized in Table 6.2.

Table 6.1 Theoretical and practical implication of this thesis

	Theoretical Implication	Practical Implication
OVERALL THESIS	- A hybrid and systematic approach of text analytics and service innovation planning methods - Contributes to the foundation of new smart service system development research	- Provides a data-driven approach embracing user innovation from outside the boundaries of the firm into their own service innovation strategies - Visualizes a concept of smart service system in the form of data-integrated technology roadmap
Chapter 3. *Planning for System Innovation*	- Proposes a quantitative approach using mobile app service documents as a remedy for subjective and expert-based morphology building	- Offers an effective way for service designer to scan the current smart service system market - Enables practitioners to capture divers and unusual elements that are to be considered as a trigger for the design of new smart service system
Chapter 4. *Planning for Technology Innovation*	- Elaborates textual CBR to extend the application area of lead user communities into an incremental or even disruptive innovation of smart service systems	- Enables service developers to foresee unexpected issues related to the smart service system and to explore more feasible solutions by leveraging lead users' knowledge
Chapter 5. *Planning for Data Innovation*	- Suggests a concept of data-integrated technology roadmap, and proposes relevant structure and typology for smart service systems	- Supports service managers to communicate efficiently with colleagues by visualizing the direction of the smart service system development

This research suggested a novel data-driven approach to the generation of new smart service system concepts, which consists of two stages. The smart service systems which have high innovative potential are identified at the first stage, based on novelty and quality measures. Novelty detection technique and sentiment analysis were employed here as a tool for dealing with mobile app service documents. Then, a number of new smart service system concepts were derived at the second stage by exploring all possible combinations of a morphology matrix. Furthermore, a case study on 'smart home' has been illustrated to show the feasibility.

The benefits of the proposed approach are as follows. First, from a theoretical perspective, this paper suggests a new and effective way to scan the current market. The existing smart service systems are classified into four categories, which are based on novelty and quality. With a Novelty-Quality map, one can easily identify what new and trendy smart service systems are being developed and released to the public. Second, from a methodological perspective, this paper proposed a quantitative approach using mobile app service documents as a remedy for subjective and expert-based morphology building. Specifically, dimensions are developed based on the characteristics of smart service systems, whereas shapes are derived from service description documents via the combination of textminig techniques. With this morphology matrix, a large number of alternatives are derived in a very short time. Third, from a practical perspective, the proposed approach provides managerial implications for the practitioners who are in charge of the new service development. It enables practitioners to capture diverse and unusual elements that are to be considered as a trigger for the development of new smart service systems. This can lead to new ideas completely different from previously existing ones. Particularly, the use of mobile app service documents provides a practical implication to the use of big data in the innovation process, by facilitating both quick decision-making and plentiful decision alternatives. Furthermore, once the overall process suggested in this study has been established, real-time monitoring is possible only by updating input data. This allows the practitioners to identify market trend through two indicators, novelty and quality, among new services that are pouring out.

Chapter 4. Planning for Technology Innovation

This research proposes a revised textual case based reasoning approach to reflect lead user's characteristics in technology-intensive product development. The utilization of lead user community is improved by separating the problem and solution SAO structures from a post shared by the lead user in case representation stage. Where the problem statement is unclear in case retrieval stage, we suggest a hybrid approach of technology tree and co-occurrence analysis, whose main focus is to identify technological structure of the product and potential problems. It enables us to explore various problems and exact solutions with expert judgement at last.

This research contributes to research in three ways. First, from the theoretical perspective, we propose a differentiated approach reflecting the characteristics of the lead user. The elaborated textual CBR extends the application area of lead user communities into an incremental or even disruptive innovation of product. Second, from the methodological perspective, this paper focuses on SAO structures instead of keywords to minimize the loss of information and employs technology tree as a remedy for providing a broad overview of issues in technology-intensive product development. Third, from the managerial perspective, this study provides an organized approach for embracing lead user innovation from outside the boundaries of the firm into their own innovation strategies. Since our approach can both efficiently identify the problems from lead users' experiences and also suggest ways to find out related solutions based on their ideas, firms can benefit by incorporating it to their R&D efforts.

Chapter 5. Planning for Data Innovation

This research suggests a concept of data-integrated technology roadmap, and proposes relevant structure, typology, and roadmapping process to show how to integrate data into the technology roadmapping. Based on practical cases of healthcare and education service industry with an extensive literature review, types of data integration are defined. Moreover, typology of data-integrated technology roadmap is suggested considering types of data integration with a newly constructed technology roadmap adding a data layer.

The contributions of this research is largely threefold from the theoretical and managerial perspective. First, types of data integration and their characteristics are established and thus make it easier to understand existing smart service systems at macro level. The interaction among the system elements – service, user, and data – also can be captured based on the three types: 'data as supporter', 'data as mediator', 'data as value generator'. Second, proposed structure of data-integrated technology roadmap visualizes the role of data for planning a certain smart service system by adding a data layer as intermediate and functional link at micro level. The inclusion of data requirements in the 'data' layer help us clarify the process of planning related hardware and software technologies. Lastly, the typology of data-integrate technology roadmap and corresponding roadmapping process is suggested briefly. It reflects not only technology-push and market-pull perspective, but also includes the direction of data-pull or data-push in terms of a new innovation driver.

Table 6.2 Detailed implication of each chapter

Chapter	Detailed Implication
Chapter 3. *Planning for System Innovation*	- Defines critical components of smart service systems based on the extensive literature review - Ensures diversity in morphology development by utilizing a larger amount of mobile app service documents - Ensures novelty and quality of system configuration by filtering high innovative potential
Chapter 4. *Planning for Technology Innovation*	- Yields insightful details of problem-solution sets from a vast amount of posts in the online lead user community - Provides ideas which can be a starting point of incremental and disruptive innovation, specifying potential problems that must be addressed in the early stage - Identifies feasible technological solutions corresponding to the retrieved problems
Chapter 5. *Planning for Data Innovation*	- Defines types of data integration in both macro (system) and micro (service) level of smart service systems - Provides analytical guidelines to understand existing smart service systems in an easier way - Visualizes data requirements and related hardware and software technologies by adding the 'data layer' as a subject of strategic planning

Despite these contributions, there still remains a substantial amount of work to be done for the development of a stronger theoretical and methodological foundation. The limitations of each study that have been presented in the sub-conclusion of each chapter can be a good starting point of future research. In addition, the overall thesis has a threefold limitation.

First and foremost, a multidisciplinary approach is required due to the inherent nature of the smart service systems. In particular, user interactions occur in a myriad of ways according to user interface (UI), human-computer interaction(HCI), sensor technologies, and network technologies. Thus, there are much things to consider besides physical hardware and software in the development of smart service systems.

Second, various smart service system concepts are designed and the potential problems while using smart service systems and related technological solutions are explored through this thesis. However, not all smart service systems can be successful in business. Thus, the evaluation process is necessary for future research to make a proper decision when targeting new smart service system among alternatives.

Third, it utilized the single source of the lead user community as an experimental study in each chapter. However, the increasing complexity of smart service systems is highlighting the role of collective intelligence. Therefore, it is necessary to develop a merge process to derive an agreed knowledge from different types of textual data sources.

Bibliography

[1] Aamodt, A. & Plaza, E. Case-based reasoning: Foundational issues, methodological variations, and system approaches, *AI communications*, 7 (1994), pp. 39-59.

[2] Akehurst, G. User generated content: the use of blogs for tourism organisations and tourism consumers, *Service business*, 3 (2009), p. 51.

[3] Albright, R. E. & Kappel, T. A. Roadmapping in the corporation, *Research-Technology Management*, 46 (2003), pp. 31-40.

[4] Allmendinger, G. & Lombreglia, R. Four strategies for the age of smart services, Harvard Business Review, 83 (2005), p. 131.

[5] An, Y., Lee, S. & Park, Y. Development of an integrated product-service roadmap with QFD: A case study on mobile communications, *International Journal of Service Industry Management*, 19 (2008), pp. 621-638.

[6] Aral, S. & Walker, D. Identifying influential and susceptible members of social networks, *Science*, 337 (2012), pp. 337-341.

[7] Arciszewski, T. Morphological analysis in inventive engineering, *Technological Forecasting and Social Change,* 126 (2018), pp. 92-101.

[8] Avramenko, Y. & Kraslawski, A. *Case Based Design: Applications in Process*

Engineering, Springer, Berlin, 2008.

[9] Balta-Ozkan, N., Boteler, B. & Amerighi, O. European smart home market development: Public views on technical and economic aspects across the United Kingdom, Germany and Italy, *Energy Research & Social Science,* 3 (2014), pp. 65-77.

[10] Barták, R., Hraško, A. & Obdržálek, D. A controller for autonomous landing of AR. Drone, in *The 26th Chinese Control and Decision Conference (2014 CCDC)*, May. 2014, pp. 329-334.

[11] Barton, D. & Court, D. Making advanced analytics work for you, Harvard Business Review, 90 (2012), pp. 78-83.

[12] Benamara, F. & Moriceau, V. Prepositions in Cooperative Question-Answering Systems: A Preliminary Analysis, In *Syntax and Semantics of Prepositions*, Springer, Dordrecht, 2006, pp. 307-330.

[13] Berry, M. W. & Kogan, J. *Text mining: applications and theory*, John Wiley & Sons, 2010.

[14] Beverungen, D., Müller, O., Matzner, M., Mendling, J. & Vom Brocke, J. Conceptualizing smart service systems, *Electronic Markets*, 29 (2019), pp. 7-18.

[15] Bogers, M., Afuah, A. & Bastian, B. Users as innovators: A review, critique, and future research directions, *Journal of Management*, 36 (2010), pp. 857-875.

[16] Breunig, M. M., Kriegel, H. P., Ng, R. T. & Sander, J. LOF: identifying density-based local outliers, in *Proceedings of The 2000 ACM SIGMOD International Conference on Management of Data*, May. 2000, pp. 93-104.

[17] Buede, D. M. & Miller, W. D. *The engineering design of systems: Models and methods*, John Wiley & Sons, 2016.

[18] Cannata, A., Gerosa, M. & Taisch, M. A Technology Roadmap on SOA for smart embedded devices: Towards intelligent systems in manufacturing, in *2008 IEEE International Conference on Industrial Engineering and Engineering Management*, December. 2008, pp. 762-767.

[19] Carrubbo, L., Bruni, R., Cavacece, Y. & Moretta Tartaglione, A. Service system platforms to improve value co-creation: insights for translational medicine, in *Service Dominant Logic, Network and Systems Theory and Service Science: Integrating three Perspectives for a New Service Agenda*, Giannini Editore, 2015, pp. 1-18

[20] Cascini, G., Fantechi, A. & Spinicci, E. Natural language processing of patents and technical documentation, in *International Workshop on Document Analysis Systems*, September. 2004, pp. 508-520.

[21] Chaitanya, T. S. K., Harika, V. S. & Prabadevi, B. A sentiment analysis approach by identifying the subject object relationship, in *2017 2nd International Conference on Communication and Electronics Systems (ICCES)*, October. 2017, pp. 62-68.

[22] Choi, S., Kang, D., Lim, J. & Kim, K. A fact-oriented ontological approach to SAO-based function modeling of patents for implementing Function-based Technology Database, *Expert Systems with Applications*, 39 (2012), pp. 9129-9140.

[23] Choi, S., Park, H., Kang, D., Lee, J. Y. & Kim, K. An SAO-based text mining

approach to building a technology tree for technology planning, *Expert Systems with Applications*, 39 (2012), pp. 11443-11455.

[24] Choi, S., Yoon, J., Kim, K., Lee, J. Y. & Kim, C. H. SAO network analysis of patents for technology trends identification: a case study of polymer electrolyte membrane technology in proton exchange membrane fuel cells, *Scientometrics,* 88 (2011), pp. 863-883.

[25] Chowdhury, S., Haftor, D. & Pashkevich, N. Smart product-service systems (Smart PSS) in industrial firms: A literature review, in *The 10th CIRP Conference on Industrial Product-Service Systems*, 2018, pp. 26-31.

[26] Christensen, K., Nørskov, S., Frederiksen, L. & Scholderer, J. In search of new product ideas: Identifying ideas in online communities by machine learning and text mining, *Creativity and Innovation Management*, 26 (2017), pp. 17-30.

[27] Cocchioni, F., Mancini, A. & Longhi, S. Autonomous navigation, landing and recharge of a quadrotor using artificial vision, in *the 2014 International Conference on Unmanned Aircraft Systems (ICUAS)*, May. 2014, pp. 418-429.

[28] Colen, P. J. & Lambrecht, M. R. Product service systems: exploring operational practices, *The Service Industries Journal*, 33 (2013), pp. 501-515.

[29] Cronin Jr, J. J., Brady, M. K. & Hult, G. T. M. Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments, *Journal of Retailing*, 76 (2000), pp. 193-218.

[30] Dahlander, L., Frederiksen, L. & Rullani, F. Online communities and open innovation, *Industry and Innovation*, 15 (2008), pp. 115-123.

[31] Duan, W., Cao, Q., Yu, Y. & Levy, S. Mining online user-generated content:

using sentiment analysis technique to study hotel service quality, in *The HICSS 2013, Proceedings of The 46th Hawaii International Conference on System Sciences*, January. 2013, pp. 3119-3128.

[32] Espíndola, D., Duarte Filho, N., Botelho, S., Carvalho, J. & Pereira, C., Internet of things to provide scalability in product-service systems, in *Proceedings of UBICOMM 2012: The Sixth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, 2012, pp. 214-221.

[33] Fantana, N. L., Riedel, T., Schlick, J., Ferber, S., Hupp, J., Miles, S., ... & Svensson, S. *IoT applications—value creation for industry,* in *Internet of things: Converging technologies for smart environments and integrated ecosystems*, River Publishers, 2013, pp. 153-206.

[34] Faraj, S., Jarvenpaa, S. L. & Majchrzak, A. Knowledge collaboration in online communities, *Organization science*, 22 (2011), pp. 1224-1239.

[35] Faria, N. & Mendes, L. Organizational image's partial mediation role between quality and users' satisfaction, *The Service Industries Journal*, 33 (2013), pp. 1275-1293.

[36] Fortunato, A., Gorgoglione, M., Messeni Petruzzelli, A. & Panniello, U. Leveraging big data for sustaining open innovation: the case of social TV, *Information Systems Management*, 34 (2017), pp. 238-249.

[37] Franke, N. & Piller, F. Value creation by toolkits for user innovation and design: The case of the watch market, *Journal of Product Innovation Management*, 21 (2004), pp. 401-415.

[38] Franke, N. & Shah, S. How communities support innovative activities: an exploration of assistance and sharing among end-users, *Research Policy*, 32 (2003), pp. 157-178.

[39] Füller, J., Jawecki, G. & Mühlbacher, H. Innovation creation by online basketball communities, *Journal of Business Research,* 60 (2007), pp. 60-71.

[40] Gambardella, A., Raasch, C. & von Hippel, E. The user innovation paradigm: impacts on markets and welfare, *Management Science,* 63 (2016), pp. 1450-1468.

[41] Geum, Y. & Park, Y. How to generate creative ideas for innovation: A hybrid approach of WordNet and morphological analysis, *Technological Forecasting and Social Change*, 111 (2016), pp. 176-187.

[42] Geum, Y., Jeon, H. & Lee, H. Developing new smart services using integrated morphological analysis: Integration of the market-pull and technology-push approach, *Service Business*, 10 (2016), pp. 531-555.

[43] Geum, Y., Lee, H., Lee, Y. & Park, Y. Development of data-driven technology roadmap considering dependency: An ARM-based technology roadmapping, *Technological Forecasting and Social Change*, 91 (2015), pp. 264-279.

[44] Geum, Y., Lee, S., Kang, D. & Park, Y. Technology roadmapping for technology-based product–service integration: A case study, *Journal of Engineering and Technology management*, 28 (2011), pp. 128-146.

[45] Geum, Y., Noh, E. & Park, Y. Generating new service ideas: the use of hybrid innovation tools to reflect functional heterogeneity of services, *R&D Management*, 46 (2016), pp. 736–748.

[46] Goh, K. Y., Heng, C. S. & Lin, Z. Social media brand community and consumer behavior: Quantifying the relative impact of user-and marketer-generated content*, Information Systems Research*, 24 (2013), pp. 88-107.

[47] Gubbi, J., Buyya, R., Marusic, S. & Palaniswami, M. Internet of Things (IoT): A vision, architectural elements, and future directions, *Future Generation Computer Systems*, 29 (2013), pp. 1645-1660.

[48] Guo, J., Wang, X., Li, Q. & Zhu, D. Subject–action–object-based morphology analysis for determining the direction of technological change, *Technological Forecasting and Social Change*, 105 (2016), pp. 27-40.

[49] Guzman, E. & Maalej, W. How do users like this feature? a fine grained sentiment analysis of app reviews, in *The Requirements Engineering Conference (RE), The 2014 IEEE 22nd International Requirements Engineering Conference*, August., 2014, pp. 153-162.

[50] Hansen, L. K., Arvidsson, A., Nielsen, F. Å., Colleoni, E. & Etter, M. Good friends, bad news-affect and virality in twitter, in *The Future Information Technology*, Springer, 2011, pp. 34-43.

[51] He, W. Improving user experience with case-based reasoning systems using text mining and Web 2.0, *Expert Systems with Applications*, 40 (2013), pp. 500-507.

[52] He, W. & Wang, F. K. Integrating a case-based reasoning shell and Web 2.0: design recommendations and insights, *World Wide Web*, 19 (2016), pp. 1231-1249.

[53] Hérissé, B., Hamel, T., Mahony, R. & Russotto, F. X. A terrain-following

control approach for a vtol unmanned aerial vehicle using average optical flow, *Autonomous Robots*, 29 (2010), pp. 381-399.

[54] Hsiao, C. H., Chang, J. J. & Tang, K. Y. Exploring the influential factors in continuance usage of mobile social Apps: Satisfaction, habit, and customer value perspectives, *Telematics and Informatics*, 33 (2016), pp. 342-355.

[55] Jang, W., Kim, J. & Park, Y. Why the online customer reviews are inconsistent? textual review vs. scoring review, in *The Digital Enterprise Design & Management*, 2014, p. 151.

[56] Jeong, Y. & Yoon, B. Development of patent roadmap based on technology roadmap by analyzing patterns of patent development, *Technovation*, 39 (2015), pp. 37-52.

[57] Jin, G., Jeong, Y. & Yoon, B. Technology-driven roadmaps for identifying new product/market opportunities: Use of text mining and quality function deployment, *Advanced Engineering Informatics,* 29 (2015), pp. 126-138.

[58] Kang, D. & Park, Y. Review-based measurement of customer satisfaction in mobile service: Sentiment analysis and VIKOR approach, *Expert Systems with Applications*, 41 (2014), pp. 1041-1050.

[59] Kim, C., Choe, S., Choi, C. & Park, Y. A systematic approach to new mobile service creation, *Expert Systems with Applications*, 35 (2008), pp. 762-771.

[60] Kim, J. & Lee, C. Novelty-focused weak signal detection in futuristic data: Assessing the rarity and paradigm unrelatedness of signals, *Technological Forecasting and Social Change*, 120 (2017), pp. 59-76.

[61] Kim, J. & Park, Y. Leveraging ideas from user innovation communities: using

text-mining and case-based reasoning, *R&D Management*, 49 (2019), pp. 155-167.

[62] Kim, J., Park, Y., Kim, C. & Lee, H. Mobile application service networks: Apple's App Store, *Service Business*, 8 (2014), pp. 1-27.

[63] Köhler, A. R. & Som, C. Risk preventative innovation strategies for emerging technologies the cases of nano-textiles and smart textiles, *Technovation*, 34 (2014), pp. 420-430.

[64] Kolodner, J. *Case-based reasoning*, Morgan Kaufmann, 2014.

[65] Kozinets, R. V., Hemetsberger, A. & Schau, H. J. The wisdom of consumer crowds: Collective innovation in the age of networked marketing, *Journal of Macromarketing*, 28 (2008), pp. 339-354.

[66] Kumar, V., Chattaraman, V., Neghina, C., Skiera, B., Aksoy, L., Buoye, A. & Henseler, J. Data-driven services marketing in a connected world, *Journal of Service Management*, 24 (2013), pp. 330-352.

[67] Kuo, Y. F., Wu, C. M. & Deng, W. J. The relationships among service quality, perceived value, customer satisfaction, and post-purchase intention in mobile value-added services, *Computers in Human Behavior*, 25 (2009), pp. 887-896.

[68] Kwon, H. & Park, Y. Proactive development of emerging technology in a socially responsible manner: Data-driven problem solving process using latent semantic analysis, *Journal of Engineering and Technology Management*, 50 (2018), pp.45-60.

[69] Kwon, H., Kim, J. & Park, Y. Applying LSA text mining technique in envisioning social impacts of emerging technologies: The case of drone

technology, *Technovation*, 60 (2017), pp. 15-28.

[70] Kwon, H., Park, Y. & Geum, Y. Toward data-driven idea generation: Application of Wikipedia to morphological analysis, *Technological Forecasting and Social Change*, 132 (2018), pp. 56-80.

[71] Larson, R. C. Commentary—Smart service systems: Bridging the silos, *Service Science*, 8 (2016), pp. 359-367.

[72] Lee, C. & Lee, H. Novelty-focussed document mapping to identify new service opportunities, *The Service Industries Journal*, 35 (2015), pp. 345-361.

[73] Lee, C., Seol, H. & Park, Y. Identifying new IT-based service concepts based on the technological strength: A text mining and morphology analysis approach, in *The FSKD 2007. Proceedings of the 4th International Conference on Fuzzy Systems and Knowledge Discovery*, August. 2007, pp. 36-40.

[74] Lee, C., Song, B. & Park, Y. Generation of new service concepts: A morphology analysis and genetic algorithm approach, *Expert Systems with Applications*, 36 (2009), pp. 12454-12460.

[75] Lee, G. & Raghu, T. S. Determinants of mobile apps' success: Evidence from the app store market, *Journal of Management Information Systems*, 31 (2014), pp. 133-170.

[76] Lee, J. H., Phaal, R. & Lee, S. H. An integrated service-device-technology roadmap for smart city development, *Technological Forecasting and Social Change*, 80 (2013), pp. 286-306.

[77] Lee, S. & Park, Y. Customization of technology roadmaps according to roadmapping purposes: Overall process and detailed modules, *Technological*

Forecasting and Social Change, 72 (2005), pp. 567-583.

[78] Lee, S., Geum, Y., Lee, H. & Park, Y. Dynamic and multidimensional measurement of product-service system (PSS) sustainability: a triple bottom line (TBL)-based system dynamics approach, *Journal of Cleaner Production*, 32 (2012), pp. 173-182.

[79] Lerch, C. & Gotsch, M. Digitalized product-service systems in manufacturing firms: A case study analysis, *Research-Technology Management*, 58 (2015), pp. 45-52.

[80] Lim, C. H., Kim, M. J., Heo, J. Y. & Kim, K. J. Design of informatics-based services in manufacturing industries: Case studies using large vehicle-related databases, *Journal of Intelligent Manufacturing*, 29 (2018), pp. 497-508.

[81] Lim, C. H., Maglio, P. P., Kim, K. J., Kim, M. J. & Kim, K. H. Toward smarter service systems through service-oriented data analytics, in *The 2016 IEEE 14th International Conference on Industrial Informatics (INDIN)*, July. 2016, pp. 936-941

[82] Lim, C. & Maglio, P. P. Data-driven understanding of smart service systems through text mining, *Service Science*, 10 (2018), pp. 154-180.

[83] Lim, C., Kim, M. J., Kim, K. H., Kim, K. J. & Maglio, P. P. Using data to advance service: Managerial issues and theoretical implications from action research, *Journal of Service Theory and Practice*, 28 (2018), pp. 99-128.

[84] Lim, S. L. & Bentley, P. J. How to be a successful app developer: Lessons from the simulation of an app ecosystem, *ACM SIGEVOlution*, 6 (2012), pp. 2-15.

[85] Lin, F. R., Hsieh, L. S. & Chuang, F. T. Discovering genres of online discussion

threads via text mining, *Computers & Education*, 52 (2009), pp. 481-495.

[86] Liu, C. Z., Au, Y. A. & Choi, H. S. Effects of freemium strategy in the mobile app market: An empirical study of google play, *Journal of Management Information Systems*, 31 (2014), pp. 326-354.

[87] Liu, D. R. & Ke, C. K. Knowledge support for problem-solving in a production process: A hybrid of knowledge discovery and case-based reasoning, *Expert Systems with Applications*, 33 (2007), pp. 147-161.

[88] Lüthje, C. Characteristics of innovating users in a consumer goods field: An empirical study of sport-related product consumers, *Technovation*, 24 (2004), pp. 683-695.

[89] Lüthje, C. & Herstatt, C. The Lead User method: an outline of empirical findings and issues for future research, *R&D Management*, 34 (2004), pp. 553-568.

[90] Maglio, P. P. & Lim, C. H. Innovation and big data in smart service systems, *Journal of Innovation Management*, 4 (2016), pp. 11-21.

[91] Maglio, P. P., Kwan, S. K. & Spohrer, J. Commentary—Toward a research agenda for human-centered service system innovation, *Service Science*, 7 (2015), pp. 1-10.

[92] Maglio, P. P., Vargo, S. L., Caswell, N. & Spohrer, J. The service system is the basic abstraction of service science, *Information Systems and e-Business Management*, 7 (2009), pp. 395-406.

[93] Marchi, G., Giachetti, C. & De Gennaro, P. Extending lead-user theory to online brand communities: The case of the community Ducati, *Technovation*,

31 (2011), pp. 350-361.

[94] Medina-Borja, A. Editorial: Smart things as service providers: a call for convergence of disciplines to build a research agenda for the service systems of the future*, Service Science*, 7 (2015), pp. 2-5.

[95] Menor, L. J., Tatikonda, M. V. & Sampson, S. E. New service development: Areas for exploitation and exploration*, Journal of Operations Management*, 20 (2002), pp. 135-157.

[96] Miorandi, D., Sicari, S., De Pellegrini, F. & Chlamtac, I. Internet of things: Vision, applications and research challenges*, Ad hoc networks,* 10 (2012), pp. 1497-1516.

[97] Mustapha, S. S. Case-based reasoning for identifying knowledge leader within online community*, Expert Systems with Applications*, 97 (2018), pp. 244-252.

[98] Nasukawa, T. & Yi, J. Sentiment analysis: Capturing favorability using natural language processing*,* in *The K-CAP 2003, Proceedings of the 2nd International Conference on Knowledge capture*, October. 2003, pp. 70-77.

[99] Ölvander, J., Lundén, B. & Gavel, H. A computerized optimization framework for the morphological matrix applied to aircraft conceptual design*, Computer-Aided Design*, 41 (2009), pp. 187-196.

[100] Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patrício, L. & Voss, C. A. Service research priorities in a rapidly changing context*, Journal of Service Research*, 18 (2015), pp. 127-159.

[101] Phaal, R., Farrukh, C. & Probert, D. *T-Plan: The fast-start to technology roadmapping: planning your route to success,* University of Cambridge,

Institute for Manufacturing, 2001.

[102] Phaal, R., Farrukh, C. & Probert, D. Customizing roadmapping, *Research-Technology Management*, 47 (2004), pp. 26-37.

[103] Probert, D. & Radnor, M. Frontier experiences from industry-academia consortia, *Research-Technology Management*, 46 (2003), pp. 27-30.

[104] Roy, R. Role of relevant lead users of mainstream product in the emergence of disruptive innovation, *Technological Forecasting and Social Change*, 129 (2018), pp. 314-322.

[105] Savransky, S. D. *Engineering of creativity: Introduction to TRIZ methodology of inventive problem solving,* CRC Press, 2000.

[106] Schatzmann, J., Schäfer, R. & Eichelbaum, F. Foresight 2.0-Definition, overview & evaluation, *European Journal of Futures Research*, 1 (2013), pp. 1-15.

[107] Shen, L., Yan, H., Fan, H., Wu, Y. & Zhang, Y. An integrated system of text mining technique and case-based reasoning (TM-CBR) for supporting green building design, *Building and Environment*, 124 (2017), pp. 388-401.

[108] Sigala, M. & Kyriakidou, O. Creativity and innovation in the service sector, *The Service Industries Journal*, 35 (2015), pp. 297-302.

[109] Spohrer, J. C. & Demirkan, H. Introduction to the smart service systems: Analytics, cognition, and innovation minitrack, in *The 2015 48th Hawaii International Conference on System Sciences*, January. 2015, pp. 1442-1442.

[110] Sureshchandar, G. S., Rajendran, C. & Anantharaman, R. N. The relationship between service quality and customer satisfaction - a factor specific approach,

Journal of Services Marketing, 16 (2002), pp. 363-379.

[111] Tan, L. & Wang, N. Future internet: The internet of things, in *The 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, August. 2010, pp. V5-376.

[112] Tuarob, S. & Tucker, C. S. Automated discovery of lead users and latent product features by mining large scale social media networks, *Journal of Mechanical Design*, 137 (2015), 071402.

[113] Valencia Cardona, A. M., Mugge, R., Schoormans, J. P. & Schifferstein, H. N. The design of smart product-service systems (PSSs): An exploration of design characteristics, *International Journal of Design*, 9 (2015), pp. 13-28.

[114] Vaughan, M. R., Seepersad, C. C. & Crawford, R. H. Creation of empathic lead users from non-users via simulated lead user experiences, in *The ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, August. 2014

[115] Verhaegen, P. A., D'hondt, J., Vandevenne, D., Dewulf, S. & Duflou, J. R. Identifying candidates for design-by-analogy, *Computers in Industry*, 62 (2011), pp. 446-459.

[116] Vermesan, O., Friess, P., Guillemin, P., Sundmaeker, H., Eisenhauer, M., Moessner, K., ... & Cousin, P. *Internet of things strategic research and innovation agenda* in *Internet of things: Converging technologies for smart environments and integrated ecosystems*, River Publishers, 2013, pp. 7-151.

[117] Vishnevskiy, K., Karasev, O., & Meissner, D. Integrated roadmaps for strategic management and planning, *Technological Forecasting and Social Change*, 110

(2016), pp. 153-166.

[118] Von Hippel, E. The dominant role of users in the scientific instrument innovation process, *Research Policy*, 5 (1976), pp. 212-239.

[119] Von Hippel, E. Democratizing innovation: The evolving phenomenon of user innovation, *Journal für Betriebswirtschaft*, 55 (2005), pp. 63-78.

[120] Von Hippel, E., De Jong, J. P. & Flowers, S. Comparing business and household sector innovation in consumer products: findings from a representative study in the United Kingdom, *Management Science*, 58 (2012), pp. 1669-1681.

[121] Von Krogh, G., Spaeth, S. & Lakhani, K. R. Community, joining, and specialization in open source software innovation: A case study, *Research Policy*, 32 (2003), pp. 1217-1241.

[122] Wamba, S. F., Akter, S., Edwards, A., Chopin, G. & Gnanzou, D. How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study, *International Journal of Production Economics*, 165 (2015), pp. 234-246.

[123] Watanabe, K. & Mochimaru, M. Expanding impacts of technology-assisted service systems through generalization: Case study of the Japanese service engineering research project, *Service Science*, 9 (2017), pp. 250-262.

[124] Weber, R. O., Ashley, K. D. & Brüninghaus, S. Textual case-based reasoning, *The Knowledge Engineering Review*, 20 (2005), pp. 255-260.

[125] Wissema, J. G. Morphological analysis: its application to a company TF investigation, *Futures*, 8 (1976), pp. 146-153.

[126] Wu, M. C., Lo, Y. F. & Hsu, S. H. A case-based reasoning approach to generating new product ideas, *The International Journal of Advanced Manufacturing Technology*, 30 (2006), pp. 166-173.

[127] Wuenderlich, N. V., Heinonen, K., Ostrom, A. L., Patricio, L., Sousa, R., Voss, C. & Lemmink, J. G. "Futurizing" smart service: implications for service researchers and managers, *Journal of Services Marketing*, 29 (2015), pp. 442-447.

[128] Xiao, X., Skitmore, M. & Hu, X. Case-based reasoning and text mining for green building decision making, *Energy Procedia*, 111 (2017), pp. 417-425.

[129] Yoon, J. & Kim, K. Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks, *Scientometrics*, 88 (2011), pp. 213-228.

[130] Yoon, J., Park, H. & Kim, K. Identifying technological competition trends for R&D planning using dynamic patent maps: SAO-based content analysis, *Scientometrics*, 94 (2013), pp. 313-331.

[131] Yu, X. & Zhang, B. Obtaining advantages from technology revolution: A patent roadmap for competition analysis and strategy planning, *Technological Forecasting and Social Change*, 145 (2019),pp. 273-283.

[132] Zec, M. & Matthes, F. Web-based software-support for collaborative morphological analysis in real-time, *Technological Forecasting and Social Change*, 126 (2018), pp. 168-181.

[133] Zhou, F., Jiao, R. J. & Linsey, J. S. Latent customer needs elicitation by use case analogical reasoning from sentiment analysis of online product reviews,

Journal of Mechanical Design, 137 (2015), 071401.

[134] Zhu, G. N., Hu, J., Qi, J., Ma, J. & Peng, Y. H. An integrated feature selection and cluster analysis techniques for case-based reasoning*, Engineering Applications of Artificial Intelligence*, 39 (2015), pp. 14-22.

[135] Zwicky, F. *Discovery, invention, research through the morphological approach,* New York, NY: Macmillan, 1969.

국문초록

많은 산업 분야에서 스마트 서비스 시스템은 데이터를 기반으로 사용자와 제품, 기술 및 서비스의 상호 작용을 통해 가치를 창출하고 있다. 완전히 새로운 데이터 소스를 활용하거나 비전통적 접근을 개발 및 적용하는 것은 기업에게 경쟁 우위를 확보 할 수 있게 한다. 자연스럽게도, 전략적 이슈에서 회사는 다음과 같은 중요한 질문에 직면하게 된다. 참신한 스마트 서비스 시스템을 어떻게 설계해야하며 사용자에게 전달할 가치는 무엇인가? 이에 대한 답은 대부분의 혁신 연구 및 사례에서와 마찬가지로 체계적인 서비스 혁신 기획이다.

이를 위해 본 학위논문은 스마트 서비스 시스템 혁신을 위한 새로운 방향을 제시한다. 구체적으로 본 논문은 스마트 서비스 시스템의 차별적 특성을 정의하고 시스템, 기술, 그리고 데이터 혁신의 측면에서 세 가지 연구 주제를 수행한다. 각 연구 주제를 해결하기 위해 다양한 방법론을 탐색하고 기존 방법론을 개선하여 적용한다. 각 방법론을 효과적으로 활용하기 위한 구체적인 프레임워크는 자세한 절차와 함께 제시된다.

첫 번째 연구는 시스템 혁신의 관점에서 모폴로지 분석을 활용하여 새로운 스마트 서비스 시스템을 설계하기 위한 구성 요소 결정에 대해 다룬다. 이 때 모바일 앱 서비스 문서를 통합하여 형태 매트릭스의 구성 과정에 객관성과 다양성을 높이는 데이터 기반 접근 방식을 제안한다. 첫째, 참신성과 품질이라는 정량적 지표를 기반으로 혁신적인 서비스를 식별하기 위한 포트폴리오 맵을 개발한다. 둘째, 새로운 스마트 서비스 개념을 창출하기 위해 전문가의 판단과

함께 모폴로지 분석을 수행한다. 제안된 접근 방법의 타당성과 효과성은 ‘스마트 홈’에 대한 사례 연구를 통해 제시된다.

두 번째 연구는 기술 혁신의 관점에서 텍스트 사례기반추론(Textual Case-Based Reasoning)을 사용하여 기존 스마트 서비스 시스템을 개선하는 방법을 다룬다. 본 연구는 SAO (Subject-Action-Object) 분석 및 기술 트리를 사례기반추론 과정에 변형하여 통합한 새로운 텍스트 사례기반추론 방법을 제안한다. 특히, SAO 구조에 초점을 맞추어 하나의 문서에서 문제와 해결책을 나타내는 SAO 구조를 분류하고, 기술 트리를 활용하여 시스템의 특정 기능과 관련된 정확한 문제 해결 세트를 탐색함으로써 점진적이고 파괴적인 혁신을 기획할 수 있도록 돕는다. 제안된 접근법의 적절성은 리드 사용자 커뮤니티를 데이터 소스로 활용하여 ‘드론(Drone)’ 의 사례를 통해 입증된다.

마지막으로, 세 번째 연구는 데이터 혁신의 관점에서 스마트 서비스 시스템을 기획하는 방법에 중점을 둔다. 본 연구는 데이터 통합 기술 로드맵의 개념을 제안하고 데이터를 기술 로드맵을 작성하는 과정에 통합하는 방법을 제시하기 위해 관련 구조, 유형 및 로드 맵핑 프로세스를 개발한다. 첫째, 스마트 서비스 시스템에서 데이터 통합의 유형은 문헌 검토 및 실제 비즈니스 사례를 기반으로 분류한다. 둘째, 데이터 통합 기술 로드맵의 개념과 구조가 제안되어 해당 스마트 서비스 시스템을 기획하기 위한 데이터 계층이 추가된다. 데이터 계층은 내부 데이터와 외부 데이터의 두 하위 계층으로 구성된다. 셋째, 데이터 통합의 유형을 고려하여 데이터 통합 기술 로드맵의 유형을 제안한다.

본 논문은 사용자 혁신과 대용량 데이터 접근을 활용하여 스마트 서비스 시스템 기획을 용이하게하는 구체적이고 체계적인 기반을 제공할 것으로 기대된다.

궁극적으로 서비스 기획자 및 개발자가 온라인 사용자 커뮤니티의 폭발적으로 증대되는 집단 지식을 활용하여 혁신적인 기회와 가능한 기술 후보를 사전에 조사할 수있는 발판을 마련 할 것으로 기대된다. 또한 제안하는 데이터 통합 기술 로드맵을 통해 서비스 관리자는 스마트 서비스 시스템의 무형적 요소인 데이터를 중심으로 시각화하여 구체적인 서비스 기획을 가능하게 할 것으로 보인다.

주요어: 스마트 서비스 시스템 개발, 서비스 혁신 기획, 사용자 혁신, 데이터 기반 접근, 데이터 통합 기술 로드맵

학번: 2014-21804