Title	タイ生命保険事業における顧客感情ナレッジマネジメント
Author(s)	SUCHARITTHAM, Nanthawadee
Citation	
Issue Date	2021-03
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/17470
Rights	
Description	Supervisor:Dam Hieu Chi, 知識科学研究科, 博士

Customer Sentiment Knowledge Management in Thai Life Insurance

Nanthawadee Sucharittham

Japan Advanced Institute of Science and Technology

Doctoral Dissertation

Customer Sentiment Knowledge Management in Thai Life Insurance

Nanthawadee Sucharittham

Supervisor: Professor Hieu Chi Dam

School of Knowledge Science
Japan Advanced Institute of Science and Technology

March 2021

Abstract

This study explores service satisfaction with Thai life insurance based on customer sentiments expressed on social media. This task provides an analytical framework of customer sentiment knowledge management that shows how to benefit from social media feedback through immediate problem-solving. After customer opinions are identified through the sentiment extraction & analysis tool, the severity of problems is prioritized.

This research presents a new social CRM that manages knowledge using a multidimensional sentiment cube to recommend processes that need to improve due to get a high volume of negative sentiment from user-generated content from social media. This method is an emerging approach that takes benefit from the sentiment cube concept with data mining. Text mining and natural language processing (NLP) is applied to extract valuable knowledge chunks with their sentiments from critiques web-blogging in the Thai language and then map each chunk to pre-defined dimensions in a cube.

In this work, a design of multidimensional sentiment cube on social CRM was demonstrated in a case study of the Thai Life Insurance Industry; dimensions are designed with consideration of standard CRM factors as well as CRM components and aspect-based sentiment analysis. Besides, we present the results of an empirical study conducted via a questionnaire survey. The concept of social customer relationship management (social CRM) is adopted by utilizing the sentiments expressed on social media as part of an active management process to improve customer satisfaction, retain customers, and recommend solutions positively received by respondents.

This study investigates service evaluation factors, claim settlement quality, policy cancellation motives, and misunderstanding problems based on demographic characteristics, life insurance attitudes, experience, and knowledge in life insurance of Thai respondents using the extracted topics from social media.

Keywords: Customer sentiment knowledge management, Multidimensional sentiment cube, Thai life insurance, Social customer relationship management (Social CRM), Text Mining

Acknowledgements

I would like to express my sincere thanks to my thesis, extraordinary advisor, Prof. Dr. Hieu Chi Dam[1,] for his invaluable help and constant encouragement throughout this research. Prof. Dr. ThanarukTheeramunkong[2,4,5] And to Prof. Dr. Ho Bao Tu[6], Prof. Dr. Mitsuru Ikeda[1], Prof. Dr. Nam Van Huynh[1], and Dr. Choochart Hareuchaiyasak[3] for their useful suggestions and kindness in examining this thesis, including reviewers who revised this work..

Besides, I am grateful for the Speech and Audio Technology Laboratory (SPT)[3] members, which provide precious Thai language sources and valuable suggestions to this thesis. Furthermore, I would like to express an additional thank you to Dam laboratory[1] members, Dr. Nguyen Thu Huong[1] and Takeyuki S. [1], KINDML members[5], JAIST friends & staff for their invaluable friendships and help to provide knowledge and programming recommendations.

Furthermore, I greatly appreciate the support to interviewees[7-13] from many organizations for providing valuable system evaluations and precious recommendations to this thesis. Especially, AIA life insurance section head, Payoongsak S. [7,10], Wipa C. [7], Sivilai P. [7] and Vantanee S. [7] and AIA Thailand staffs[7] who provided life insurance knowledge to this thesis including section head of the other companies, Vichai J[9], Naratip J[11], Jaturong K[12], and Rattawadee A. [8] for providing service knowledge and questionnaire survey.

Finally, I most gratefully acknowledge my parents and my friends for all their support throughout the period of this research.

1. Japan Advance Institute Science and Technology (JAIST)
2. Sirindhorn International Institute of Technology (SIIT)
3. National Electronics and Computer Technology Center
4. Thammasat University (TU)
5. Knowledge, Information, and Data Management Laboratory
6. Vietnam National University
7. American International Assurance (AIA)
8. Thai Cardif Life Insurance
9. C.P. All Public company
10.Samart Telecommunication
11.Alliance for Supporting Industries Association (A.S.I.A.),
12. GetSmart service
13. Total Access Communication (DTAC)

Table of Contents

This dissertation was prepared according to the curriculum for the Collaborative Education Program organized by Japan Advanced Institute of Science and Technology and Sirindhorn International Institute of Technology (SIIT), Thammasat University.

List of Tables

List of Figures

Chapter 1
Introduction

1.1 Significant and Problem

According to the Swiss Re Institute's sigma research, the global insurance market will experience aggregate emerging-market growth of around 4.9 percent annually between 2019 and 2020, following 4.7 percent growth this year (2018). Thailand is an emerging market with continuing growth in the life insurance market. In 2016, Thai life insurance ranked ninth among the top 15 emerging marketing of the 40 largest markets by life premium volume.

According to the Thai Life Assurance Association (TLAA), Thailand's life insurance status at the end of 2017 was as follows. The number of inforce policies was 7,939,652. The number of new active policies was 1,292,748 (up 16.28%); however, there was a loss in volume of current inforce policies of 272,117 surrendered policies and 485,393 lapsed policies (a drop of 9.54%). Based on this, we found several encouraging signs for the life insurance industry in Thailand. However, the insurance market also faces the obstacle of policy loss through cancellations and lapses. The high volume of policy loss shows the significance that there are hidden problems that require attention. One way to reduce policy losses is to increase customers' satisfaction with their policies.

Customer satisfaction is an indicator of an organization’s success, mainly when it includes the service quality dimension. Therefore, an organization can use customer satisfaction to manage and develop its business [71], increasing the persistence rate for life insurance. Nonetheless, it is not easy to increase customer satisfaction in this industry. Life insurance is an intangible service, and its lack of physical dimensions makes it difficult to measure service satisfaction. The service characteristics are intangible and not readily defined. Moreover, the service delivery process is targeted at customer satisfaction and not at a physical good. Furthermore, the service is heterogeneous, as it has no exact service standard. It also depends on the quality of each personal contact [21]. These reasons have motivated many researchers to study customer satisfaction in an intangible service industry, such as life insurance.

Previous studies on customer satisfaction in the life insurance industry have often focused on identifying important indicators that impact customer satisfaction in a specific area. For example, Subashini and Velmurugan (2016) studied policyholder satisfaction in Coimbatore District, India [78]. Kannan (2018) also studied customer satisfaction in Chennai, India [40]. Some scholars have had a specific purpose or objective, such as Kuhlemeyer and Allen (1999), who explored consumer satisfaction in life insurance products purchasing (Kuhlemeyer & Allen, January 1999). Nguyen et al. (2018) presented the determinants for customer satisfaction and loyalty in life insurance services in Vietnam, and so on [43], [59]. Previous studies have focused on the passive management of problems, using questionnaires that include questions framed by service-provider experts constituting knowledge from the inside-out. For the active management of problems, the questions need to be based on outside-in knowledge by identifying problems from customer criticisms expressed externally, for example, on the internet. Resolving these problems before they become severe is essential. Therefore this research examines the severity of service dissatisfaction based on customer criticisms expressed on the internet. It uses the concept of social customer relationship management (Social CRM) to identify high-severity problems. In general, the goal is to increase customer satisfaction by resolving the main problems that cause the most dissatisfaction.

CRM is traditional strategic method companies use to manage their relationship with customers [79]. However, the emergence of social media has challenged CRM's traditional concept (Weinberg et al., 2011). Traditional CRM assumes that customers are passive, responding to a company's action primarily through purchasing behavior (Malthouse et al., 2013). The new concept of social CRM considers customers as active participants [46], [70], [74]. Social CRM (SCRM) means building relationships with customers based on social interaction. It extends the traditional customer relationship management concept by helping companies communicate better and create more effective relationships with customers, thereby enabling them to respond online. In this context, communication and information sharing occur quickly, and companies respond to customers' needs more accurately. For example, when a customer generates content on social media related to a specific brand, they are engaging with that company as an active customer.

In the context of Thai life insurance, sharing experiences on a product or service from social media often be blown out of proportion. When customers experience poor service, they usually prefer to share this experience with their friends and social media rather than to submit complaints directly to the service provider [46]. These negative messages affect the attitudes and opinions of other customers. Thai review blogs, such as Pantip.com, Krapook.com, which are free sites, contain numerous posts on service dissatisfaction, leading to negative word of mouth.

Knowledge management on customer sentiment, such as utilizing the sentiments expressed on social media to improve customer satisfaction as an active management concept of social customer relationship management (social CRM) is essential. The customer sentiment extracted from social media can be an initial source for sharing an understanding of customer behavior and finding out the improvement methods in order to increase customer satisfaction as co-creation new perspective of problem-solving. Especially, negative sentiment can utilize for finding problem-solving methods by sharing those sentiments to survey their opinions from various demographics customers based on the problems related to that negative sentiment using a questionnaire.

However, sharing an extracted customer sentiment from social media between the customer who cannot identify the characteristic of customers (as anonymous source) and customers who can identify qualifications and demographics in order to get recommendations about that sentiment is not easy. To propose the methodology to utilize the knowledge from customer sentiment for improving customer satisfaction in Thai life insurance, knowledge sharing process is a key important to merge knowledge between outside-in and inside-out can reveal customers’ mind. This process requires the critical key elements. It consists of business knowledge understanding, analysis concept understanding, and sentiment extraction technology, and sentiment analysis concept. This study chooses the Thai life insurance industry to be a case study, one of the extensive services and impacts on Thai society.

To identify service-dissatisfaction issues related to insurance from such web-blogs requires knowledge on extracting customer opinions requires technology such as natural language processing (NLP), text mining, or sentiment analysis (Liu, May 2012), (Fangtao et al., 2010), (Minqing & Bing, 2004) [10, 31, 32, 45]. It can help develop a tool that extracts customer sentiments from the text by identifying the negative words related to each aspects of Thai life insurance service. This research developed a sentiment extraction tool and analysis tool named multidimensional sentiment cube [60, 61]. The multidimensional sentiment cube results show that the customer dissatisfaction scale prioritizes the crucial problems with sentiment scoring. The causes of dissatisfaction in the claim settlement process, policy cancellation motives, and misunderstanding between policyholders and service providers are provided.

In summary, this research has purpose to present the methodology for utilizing customer sentiment to improve customer satisfaction in life insurance service using technology and knowledge sharing. Out analysis tools “Multidimensional sentiment analysis” can analyze the existing problems from negative customers’ sentiment extracted of web-sites, next synthesis analysis results to be knowledge from customer sentiment. Then, make knowledge sharing process by conducting the questionnaire to find out appropriate problem-solving methods to deal with dissatisfaction, The results are from respondents who have different qualifications based on demographics such as gender, age, income, and education and hold different life insurance, experience, knowledge, and attitudes. Finally, this research reveal the suggestions to life insurance service provider by providing practical recommendation on current problems and create new knowledge to get better customer satisfaction.

1.2 Research Problems

1. Life insurance is the sound-reducing the effect of unexpected vital accidents by sharing risk among a group of insurance members. Most people usually do not realize its importance, its benefit. Moreover, life insurance knowledge is complicated and challenging to understand the rules and regulations clearly. With this, a reason makes Thai life insurance does not grow so fast. Also, life insurance is one intangible product. It is challenging to realize the use of that, so it is essential to clearly understand the customer in dissatisfaction problems by making use of customer sentiment. Especially, service industry such as life insurance has intangible characteristics because of it is difficult to identify aspect such as product. For example, camera has main aspect on lens, weight, price, etc. While to identify service aspect require knowledge such as business process knowledge or service knowledge.

2. While turning to the Internet society, many industrials require adaptive management. To take into account under social CRM is the one useful concept for adaptation of business. Utilizing user-generated content for sharing knowledge with customers can present customer satisfaction levels and deliver a new perspective using customer sentiment knowledge management. Sharing knowledge inside-out information (passive) and outside-in information (active) require customer sentiment knowledge management.

3. In social media era, increasing efficiency of information analysis requires technical knowledge related to sentiment analysis. Making an application becomes more practical for real usage in sentiment extraction from social media is necessary by concerning the relationship among multiple business factors such as a hierarchical view based on negative, positive, or neutral. Good visualization, which presents the analysis viewpoints among multi-factors is required. Since the relationship is across different features, it reveals more meaningful analysis and can reveal hidden problems by getting the

relationship's negative opinion. For example, the A. feature represents a hidden problem when getting a negative sentiment score from A_1, which is under A. If A_1 has high-frequency co-occurrence in a negative score, it shows A_1 can present A. is a significant problem as well.

1.3 Purpose of Study

To propose the methodology to utilize customer sentiment knowledge for improving customer satisfaction in Thai life insurance

1.4 Research Questions

Main Research Question: How to utilize the knowledge on customer sentiment for improving customer satisfaction in Thai life insurance?

Sub Research Question 1: How to develop analysis tool for extracting customer sentiment knowledge from social media?

Sub Research Question 2: What are the issues relate to customer sentiment in Thai life insurance from knowledge sharing process?

Sub Research Question 3: What kinds of knowledge on customer sentiment have been utilize for knowledge sharing process between service provider and customer?

1.5 Significant of the study

Academic Significance:
To present the methodology to utilize customer sentiment in order to improve customer satisfaction in Thai life insurance service using technology and knowledge sharing.

Social Significance:
To reveal a suggestion to life insurance service providers by providing practical recommendations on current problems to get better customer satisfaction

1.6 Contributions of the study

To propose the methodology for utilizing customer sentiment from social media to improve customer satisfaction in Thai life insurance service by knowledge sharing between outside-in and inside-out knowledge

To propose the methodology in an analysis tools in sentiment analysis called multidimensional sentiment cube by deploying the benefit from OLAP. It make this tools efficiency to analysis unlimited viewpoints which more high efficiency for sentiment analysis than other tools such as single or hierarchical aspect which limited view.

1.7 Thesis Structure

The rest of this thesis structured is as following:

1.7.1 Chapter 2

This chapter is a literature review of a related research topic and previous researchers' previous work. The details are previous and current works from many researchers related to our research in this thesis. The corresponding theory consists of knowledge management, customer knowledge management, Thai life insurance, traditional customer relationship management (CRM) & social customer relationship management (social CRM), life insurance knowledge extraction with data mining, sentiment extraction for life insurance, technology related to sentiment extraction in Thai language (text mining, natural language processing (NLP)), a hierarchical concept in online analytical processing (OLAP), and multidimensional cube.

1.7.2 Chapter 3

This chapter shows the background and idea of implementing customer sentiment knowledge management and its process in Thai life insurance. The concept is explained as a knowledge management process, including the implementation of support tools; the multidimensional sentiment cube design is described and illustrates detailed implementation, including information extraction and sentiment calculation. Also, the sentiment knowledge sharing using a questionnaire is explained.

1.7.3 Chapter 4

Their research results are explained in this chapter, including examples of multidimensional sentiment cube analysis scenarios and domain experts and end-users' evaluation. In addition, the results of customer sentiment knowledge sharing using questionnaires are included.

1.7.4 Chapter 5

Finally, discussion and conclusions are made in this chapter. It consists of multidimensional sentiment cube comparison to previous works, conclusion, and recommendation from multidimensional sentiment cube analysis results. Also, an analytical framework of customer sentiment is in this chapter.

Chapter 2

Theory and Literature Review

This chapter discusses the relevant theories and investigates the previous and current researches related to this research thesis. Literature review gives explanations about the corresponding research about the knowledge management, customer knowledge management, Thai life insurance, traditional customer relationship management (CRM) & social customer relationship management (social CRM), life insurance knowledge extraction with data mining, sentiment extraction for life insurance, technology related to sentiment extraction in Thai language (text mining, natural language processing (NLP)), a hierarchical concept in online analytical processing (OLAP), and multidimensional cube. The previous research about customer satisfaction in life insurance service is also reviewed. Many previous works are described and analyzed to get more information and understanding for applying to our method.

2.1 Knowledge Management

Knowledge Management, also known as KM, is a management approach for operating within the organization to make definitions on corporate knowledge, including collecting, creating, and distributing knowledge of the organization throughout the organization to build up the knowledge and apply knowledge to gain benefits. It also creates a culture of learning within the organization, namely the collection of existing knowledge scattered in the person or document. To develop a system, everyone in the organization can access knowledge and improve themselves as an educator. Moreover, apply the knowledge to make the operation more effective.

Knowledge comes from education, research, or experience, including practical ability and understanding skills or information obtained from experience received from hearing, listening to thoughts, or practices in each area. *(Source: Royal Institute of Thailand (RIT).*

Knowledge is an integration of experience, values, contextual information, and expert insights with a framework for evaluating and integrating new experiences and data. It arises and is applied in the minds of knowledgeable people. It is frequently embedded in documents or repositories and organizational routines, processes, practices, and norms in an organization. (*Working Knowledge: How Organizations Manage What They Know).*

Knowledge Type:

Explicit knowledge is knowledge derived from theories, definitions, or manuals that anyone can access or learn. It can be transmitted through various methods such as written notes, manual theories, documents, rules, and procedures for media operations such as VCD DVD Internet. It also can be called “Concrete knowledge.”

Tacit knowledge is the knowledge inside of individual persons, such as experience, talent, instinct, or flair of each person to understand the work or information. This knowledge cannot easily be conveyed through letters. Therefore it sometimes can be called "Abstract knowledge," such as work skills, crafts, experiences, ideas.

Elements of Knowledge Management:

Knowledge management relies on three main components as follows:

- **People**, people or personnel are the most important factors in knowledge management implementation because people are the most important source of knowledge and people who will bring knowledge through organizational development.
- **Technology** is a search-exchange assistant. The key is to store data for enabling people to utilize various kind of knowledge into organizational development.
- **Knowledge process** Is management in every Steps of knowledge management In every part Must be related in a balanced way Including efficiency

The concept of knowledge management:

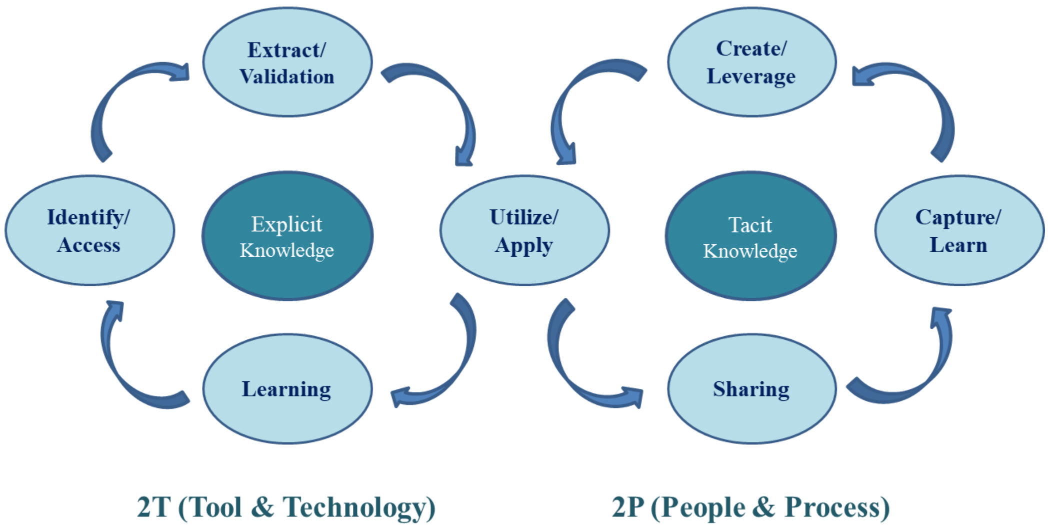


Figure 1: The concept of knowledge management

1. Identify and Access knowledge: to consider what the organization has a vision, mission, strategy, and goal to achieve the goal. What needs to be used? What and where is the current knowledge? And from whom?
2. Extract / Validation: understanding, synthesis and validation.
3. Utilize / Analysis Results: for improving, modifying, or creating some knowledge to be suitable for organization usage.
4. Utilize / Analysis Results: for improving, modifying, or creating some knowledge to be suitable for organization usage.
5. Capture and Learning: make use of knowledge to support making decisions, problem-solving, and make a part of the work, such as learning systems from knowledge creation.
6. Capture and Learning: make use of knowledge to support making decisions, problem-solving, and make a part of the work, such as learning systems from knowledge creation.
7. Implementation and Learning: to make continue to circulate continuously.

2.2 Customer Knowledge Management (CKM)

Customer Knowledge is one of a strategic resource in a company success. CKM is a necessary domain of Knowledge Management (KM) and it combines the connection between KM and Customer relationship management (CRM). CKM is an management principle which KM is an instrument to support the knowledge management between customer and organizations. It uses for manage customer relationship in order to improve CRM process such as customer acquisition, customer retention etc.

CKM focuses on defining customer knowledge, sources and types. There are some CRM previous works. For example, Gebert suggested the CKM model on four knowledge aspects: 1) content, 2) competence, 3) collaboration, and 4) composition. Rollins et al. stated CRM is a learning process where organizations make process of knowledge exchange in efficient way. They has purpose to make understanding on an integrated management approach of CKM.

Previous works example of customer knowledge management, Rollins & Halinen (2005) gave an understanding of the customer knowledge management that it requires the way to integrate a tentative theoretical framework as an integrated management approach. Jay (1990) explained the six necessary ingredients for knowledge management which will increase a chance to success of company. Ranjit & Vijayan (2003) suggested the CRM framework through the knowledge management technology which can be basis of CRM improvement. Wu, Guo and Shi (2013) provided a conceptual framework to explore the linking mechanisms between IT-based on business and customer knowledge management for model innovation including integrated both customer perspective and firm perspective for helping company understand the linking mechanisms and gain more benefits.

2.3 Life Insurance definition

The definition of life insurance in this study

Life insurance is a contract between an insurer and a policyholder in which the insurer guarantees payment of a death benefit to named beneficiaries upon the insured's death. The insurance company promises a death benefit in consideration of the payment of premium by the insured. The purpose of life insurance is to provide financial protection to surviving dependents after the death of an insured.

Life Insurance Basic is the primary insurance that will be the insurance that protects the insurer in two ways: when living or die. If the insured lives up to the last day of the contract, the insured will receive insurance plus additional benefits (if any), but if the insured dies before the last day of the contract, the company will pay (the sum insured) to the beneficiary, which the insured can specify to give to anyone.

Life Insurance Riders
Many insurance companies offer policyholders the option to customize their policies to accommodate their personal needs. Riders are the most common way a policyholder may modify their plan. There are many riders, but availability depends on the provider.

Life Insurance Company / Service provider is a company licensed by the Minister of Finance to operate a life insurance business to guarantee loss or damage to individuals or groups of people by promising to pay compensation to the insured or beneficiary in case the insured has died including other coverages as contract.

Life Insurance Agent is a person who is assigned to persuade the person to make an insurance contract. An insurance agent may be an employee of an insurer or a third party.

Insured is a person who is protected by insurance policies in life insurance.

Premium payments are set using actuarially based statistics. The insurer will determine the cost of insurance (COI), or the amount required to cover mortality costs, administrative fees, and other policy maintenance fees. Other factors that influence the premium are the insured's age, medical history, occupational hazards, and personal risk propensity. The insurer will remain obligated to pay the death benefit if premiums are submitted as required. With term policies, the premium amount includes the cost of insurance (COI). For permanent or universal policies, the premium amount consists of the COI and a cash value amount.

The death benefit is the amount of money the insurance company guarantees to the beneficiaries identified in the policy upon the insured's death. The insured will choose their desired death benefit amount based on the estimated future needs of surviving heirs. The insurance company will determine whether there is an insurable interest and if the insured qualifies for the coverage based on the company's underwriting requirements.

Hospital and Medical Expenses insurance is one type of insurance that protects medical expenses.

Policy Lapsed is an expired or terminated policy when the insured does not pay the premium at the time specified in the policy contract.

Policy Maturity is the end of the life insurance contract, which the insurer had paid total premium according to the insurance contract.

Premium is the amount paid to the life insurance company to purchase the coverage by paying in installments until the end of the contract (As agreed in the payment terms)

Sum Assured is the amount specified in the policy, defined as benefits, which must be paid when an insured event occurs.

2.4 Traditional CRM and Social CRM

Most companies have many tangible assets to support their businesses, such as knowledge workers and business premises. However, the greatest asset that a company has is the customer, as expressed by Compbell and Cunnigham (1983) [41].

As the most significant asset for a company, customer satisfaction is the main target for every company to achieve its business goal. To increase customer satisfaction, which is the key to success in business, customer relationship management

(CRM) is applied as a strategic tool for a company to build, manage and strengthen the relationship with its customers [15] [47] [58].

Customer Relationship Management (CRM) definition is

a. Develop and maintain a mutually beneficial long-term relationship between the organization and its target customers with business strategy by concerned customer-centric.

b. It is an organization's strategy that tries to reach customer feelings, understand customer behavior, and persuade a customer by creating customer relationships for customer acquisition, customer retention, and royalty, including increasing customer profitability.

Customer Relationship Management (CRM) forms is including.

1. CRM Functional and operational means process and technology which deploy to provide the best service to an organization's customer for increasing efficiency and accuracy in the operation process of organization
2. CRM customer segmentation uses customer lifetime value information and provides service in a specific group of a customer by understanding customer behavior in time. CRM can provide a fashionable product, improve service and product by innovation as customers' needs have changed fast.

As the era of social media, there are increasing the challenging of the traditional CRM framework to manage the substantial information of the customer to manage relationship among them. (Payne and Frow 2005; Verhoef, Venkatesan, et al. 2010)[47]. A typical CRM usually performs a statistical analysis of sales/service records and customer feedback to reveal customer participation patterns or behavior.

With the fast growth of ICT, both sales/service records and customer feedback are digitalized and even acquired online via a web browser or social media. The client is no longer limited a passive role in the relationship with the company. [47]. Moreover, the competitive of product will increase in various media such as mobile devices. It is trend to increasingly difficult to manage the product/service information to customer in order to gain better percept. (e.g., Schultz, Malthouse, and Pick 2012) [47].

> Harvard Business School's view on customer loyalty research about *"Customer referrals, endorsements and spreading the word"* has significance for customer behavior, impacting an organization, products, or service by "Word of mouth," especially in the customer acquisition stage. For example, "Would you like to recommend other people to buy or request service from our company?" The company can get the correct answer from a customer more than ask directly to the customer that it will return to use our service or not. These are significant issues that are ignored by many organizations now.

For this reason, many researchers have begun to study the social CRM which emerges along the border between CRM and social media. Researchers are engaged in social media impact on traditional CRM problems such as customer churn (Nitzan and Libai (2011) and Haenleim (2013) [47]. Much research is related to the consequence of word-of-mouth acquires customers compare to other channels (Trusov, Bucklin, and Pauwels 2009; Villanueva, Yoo, and Hanssens 2008). Furthermore, the result of such research found that the viewpoint of customer value has changed. The customer value is no longer limited to purchase-based and includes social elements such as customer influence, referrals, and knowledge (Kumar et al.2010; Weinberg a dBerger 2011).

Thus, there is some research about the design of the social CRM framework, which examines how CRM needs to adapt to the rise of social media to understand how social media affect CRM. They identified the company's customer engagement level on social media, which is affected by its approach to three components of traditional CRM. It consists of Acquisition, Retention, and Termination [47].

Customer acquisition: Companies seek to enlist social media in their customer acquisition efforts begin by uploading advertising spots on social media such as YouTube, Facebook, etc. This action can help companies create awareness and change attitudes among prospective customers, thereby contributing to new customers' acquisition [47].

Customer retention: Company can incorporate social media to retain existing customers and maintain an ongoing relationship with customers using social media to promote new campaigns for current customers [47]. Social media's viral benefits for customer retention are finding out the problems that scatter on critiques on various media. Usually, people will not return to the back company to give feedback directly if they do not face any severe problems and request help from the company. At this point, it makes the company miss important dissatisfaction information from a customer. It may have significant impacts on business growth in the future. However, people are familiar with expressing feeling both negative and positive on social media. It is one of the company's opportunities to do an active process like promotions on social media and do a passive process by finding out information such as the dissatisfaction feeling of a customer on social media.

Customer termination: A customer termination case is possible to initiate either by the customer or by the company. Traditional CRM can use historical records of customer activities/behavior to find out a pattern for predicting churn. All of the terminated customers' records come from the database, which key-in/input from the survey process or complain system. Both of process is a passive process after terminated problems have occurred already. The answer to each problem/or dissatisfaction problem was led by a human/expert who only makes a questionnaire.

Currently, social CRM can find out more benefits from finding a real cause of termination. The company can do active action by finding out the dissatisfaction problems that terminated customer posted on social media freely and can find more hidden problems out of the frame of expertise in this area. However, the limitation of finding information from social media is that we cannot find the real cause of negative sentiment in every terminated case. It depends on terminated customers express feeling on social media or not, and our system can track these problems.

Social media influences customer engagement in company approaches composed of customer acquisition, customer intention, and customer termination in the social CRM conceptualization. This concept is derived by analyzing the large quantities of data from customers' critiques records posted on social media. Information technology can search for knowledge by using various data platforms, including unstructured data from social media. After knowledge is extracted, the next step of the CRM process/step is to set up a strong strategy by management level and strategic expertise. Strategic preparation is the step of making the market of this strategic and follow-up and feedback process also need to set up in advance for the outstanding performance of results. In a part of strategic information, social CRM is highly concerned about the sentiment on social media. Many customers feel free to express an opinion on various websites that cannot feedback to the company directly are scattered. Businesses can gain benefits from this opportunity from customers' sentiment analysis.

In the concept of social CRM, it is essential to study CRM's current practice in the business field. How technology, especially information technology, can increase the ability of CRM to a new generation of social CRM based on customers' opinionated comments from a particular website (41). However, today's CRM system cannot benefit from this massive amount of information because it focuses on the purchase pattern in historical data of the CRM database, which keeps data records from daily company records. To enhance the CRM system, emerging the components of technology and sentiment analysis, people's opinion is necessary. This task used social CRM under traditional concepts to declare features (aspects) for sentiment analysis using techniques such as data mining, text mining, and multidimensional database for presenting analysis viewpoints of the orientation of customer's opinions regarding a particular product or service.

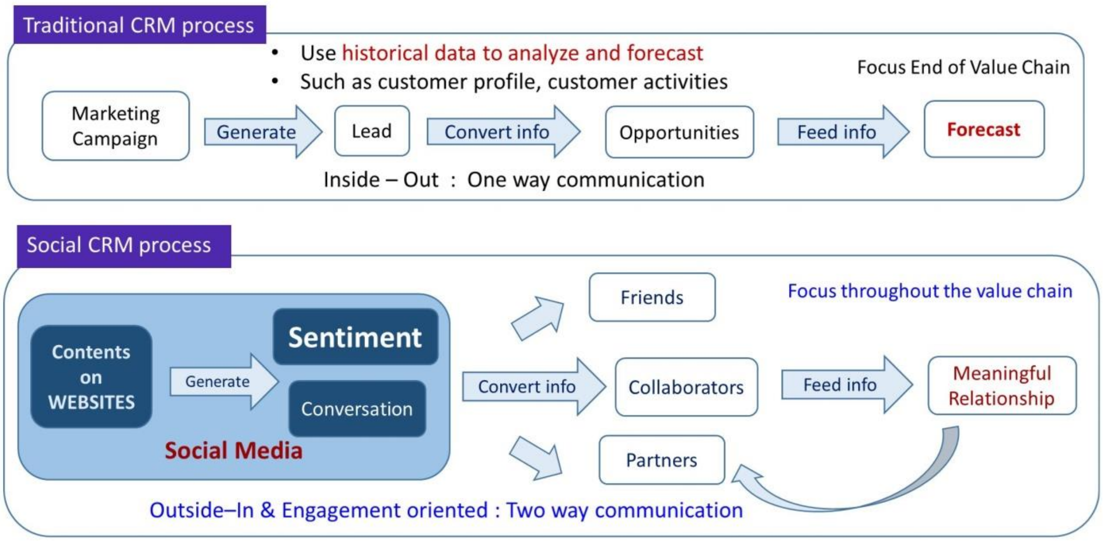


Figure 2: The different between TCRM and SCRM

There are differences between the Traditional CRM process (TCRM) and the Social CRM process (SCRM). TCRM focuses at the end of the value chain, such as the forecast process, using historical data such as customer profile and customer activities. For example, to utilize data from the current marketing campaign to generate lead of campaign then convert into opportunities and feed into forecast process to predict. Besides, TCRM is one-way communication as only an inside-out management approach. SCRM has concerned with a meaningful relationship. In the social media era, the various contents on websites provide customer sentiment on conversations. Those kinds of content can spread widely on social media and impact customer opinion when making decisions on products or services. SCRM is two-way communication between a customer and business owner as of the inside-out and outside-in management approach.

2.5 Life Insurance Knowledge Extraction with Data Mining technique

Life insurance business requires information for making a decision on many functions, and currently, they deploy various technologies to support information extraction and analysis such as a statistic, data mining, text mining, etc. This section describes the examples of data mining techniques applied to extract knowledge from life insurance. The survey covers data sources, mined knowledge, and mining techniques.

As a classic service but complex structures, insurance becomes an interesting target field of study for service science. Currently, due to various conditions and needs, a large number of insurance types and policies are developed. With a suitable type and policy, the customer and/or his/her family can sufficiently obtain income when an unexpected inconvenient event, such as an accident, sickness, or fatal death, takes place. Looking as a whole, good insurance plans can reduce the risk of both customers and insurance companies. However, life insurance is also not so popular in most developing countries, including Thailand, due to the difficulty of insurance terminologies, the complexity of insurance policies, and the intricacy of benefit assessment. Consulting proficiency is required to extract a customer's behavior and life-style from massive data to describe each policy's effect on the customer and determine the level of coverage for individual people [1]. Recently information technology has been developed to the level of practice.

Knowledge utilization provides an insightful understanding for people and organizations, such as allowing an operation by their demands, making coordination between the parties, or facilitating information management to make better decisions. Knowledge extraction is the process of exploring and constructs knowledge from various sources of data. This knowledge is extracted into solid results for easy understanding, and we can utilize this useful knowledge to support our purposes. One of the valuable data sources in a company is its database storing the customer activity records. It is possible to analyze customer behavior from such a database, including both structured and unstructured data. For instance, the structured data might consist of demographic information, including private information. An unstructured data is narrative information such as opinions from customers and evaluation results of underwriting. By utilizing knowledge extraction from these structured and unstructured data in the insurance industry to analyze customer behavior, we can understand how to obtain and dispose of products. A new customer can consume this knowledge acquired from the current customer behavior if there are the same behavior, lifestyle, and conditions. However, it is not straightforward to figure out which insurance policy can be recommended to a specific customer. Even if the most similar customers, they may occupy different minor characteristics, lifestyles, and conditions. Customer behavior can be exploited by using knowledge extraction to support various purposes from insurance activities. The initial criteria which can determine the basic rules for assigning an insurance type to an individual customer have three viewpoints.

2.5.1 Customer segmentation extraction

The importance of customer segmentation is to make new customers know their current positions, which segment matches up character or lifestyle. Because the customer separations are used for separating the group by determining the similarity of members, they make more understand each customer segment's unique characteristics.

Generally, the insurance company gets demographic data of customers when a customer made a new contract. They can utilize customers' demographic data such as education, gender, income level, marital status, age, premium, and life stage for extracting customer segmentation. A data mining technique is called It is a "clustering technique" which tries to group a given data into a class so that the members within the cluster are highly similar in comparison to one another. But these members are very different from other clusters. Similarities or differences are evaluated based on attribute values describing the element using distance measurement. Generally, there are many of clustering techniques such as hierarchy method, partitioning method, K-mean clustering method, density-based method. However, we give an example of how the clustering technique works in customer segmentation from K-means clustering.

Customer segmentation can perform extraction using clustering techniques based on determined attribute values to create new catalogs. The expected result from this analysis target is to find out the position of newcomers in the current insurance market. A new customer will know which segmentation matches his character or his lifestyle by demographic customer data analysis.

Refer to Table 1; there are provided attributes from history demographic customer data such as gender, age, occupation, income. An analysis attribute is the life stage attribute. Customers who pay attention to making an insurance contract should know their position from the results after exercise clustering techniques.

Gender	Age	Occupation	Income	Life Stage
Male	7	Student	0	Juveniles
Male	21	Soldier	15,000	First Jobbers
Male	61	No job	15,000	Retirees
Female	14	Student	0	Juveniles
Female	35	Officer	35,000	Matured
Female	55	Teacher	65,000	Matured

Table 1: Sample history demographic customer data

The data from the determined attributes can perform clustering to separate the cluster of the customer. In brief, the steps of K-means consist of do partition objects into non-empty subsets, compute the mean point of the clusters of the current partition, assign each cluster center to the mean of its assigned item and then repeat until convergence. Since clusters are not predefined, a domain expert is often required to interpret the created clusters' meaning. After taking an experiment from K-means clustering, the supposed results of a first cluster are male, 32, soldier, 16,500, first jobbers, and a second cluster is male, 35, officer, 25,000, matures in terms of gender, age, occupation, income, and life stage. Next, we put test data as male, age is 30, occupation is officer, and income is 20,000. In this case, the cluster of this test data is cluster one, and the life stage should be first jobbers. It might be different from finding out a result from fundamental analysis, which may classify this case to be matured.

Many researchers often utilized clustering techniques in their tasks. Various papers used K-means algorithms in their experiments, such as A.B. Devale and Dr. R.V.Kulkarni used the K-means clustering technique to classify a customer's group from demographic information such as gender, age, income, education level and occupation [72]. This section introduced the basic concept of each data mining concept for knowledge discovery in the insurance business. In clustering, they used K-Means to give an example if a company does not have any predefined labels. They presented based on the grouping outcome; they will target marketing and advertising campaigns to the different groups for particular types of policy. Shu-Hsien Liao et al. utilized the concept of data mining in real work to determine the level of customers' requirements in life insurance products by generating the knowledge rules and patterns from their demand chains of customers [98]. This paper also used K-means clustering to classify the customer group from demographic data, which is quite different. They used necessary information such as name, gender, birth date, and life stage needs, residential house level, life partner, and transportation mode. Depth details of features can distinguish the classification of customer groups more clearly.

Many previous works classified customer groups for different purposes from the clustering method. Not only the clustering method but also other methods can utilize for customer segmentation. For instance, the association rules are widely utilized, especially market baskets theory for finding existing customer patterns such as Yu Yan and Haiying Xie recommended using association analysis in the customer behavior model to improve the value of a customer in the customer relationship management task (CRM) of insurance industry [97]. For example, the recommendations of new insurance service to an old customer to upgrade sales volume and improve the return rate of sale-investment. Moreover, they also gave some recommendations about a basic level of informatization in pattern construction. Yongqiang Chen & Leifang Hu employed data mining knowledge into customer relationship management (CRM) in the insurance business to enhance the customer value model and market classification model [49].

2.5.2 Occurrence product patterns extraction

Finding out a suitable life insurance product for each individual is difficult because various kinds of insurance products are basic product types and rider concurrence products. They are challenging to understand the usefulness of each product's different purposes and difficult to know which group of a product should be held simultaneously to support the highest benefits.

- At present, selling channels or agents still offer a new product to a customer from their experience without considering the highest benefit of a customer. Because it is not straightforward to understand clearly the relationship of product benefits between basic and rider, so they cannot recommend the right product to the right person correctly. Usually, they offer base on their idea and sometimes base on their commission of each product at that time. Usually, life insurance product has two product types. The first type is called a basic product, which means one main policy has only a basic product for each insured. The second type is called a rider product, which is a supplement product into a basic product. Each basic product has a different purpose. Some of the basic product has a purpose of building up cash value, and others are a protection purpose in the short term. The example of a basic product in different purposes such as [67]
- Term life insurance is for a limited time, such as 10, 15, or 20 years.
- Universal life insurance is flexible and allows increasing premium over time
- Whole life insurance is permanent insurance and provides lifetime coverage. The premiums are fixed and build cash value - the value functions as a saving account that may be tax-deferred.

Rider product is the additional benefits that can be added into a basic product. Rider product has increased a level of coverage and can be blended, for an additional cost, according to future coverage needs [38]. Here are some examples of standard riders in life insurance.

- Waiver of Premium (WP) is a rider attached to the basic policy to maintain coverage in the event of total and permanent disability; subsequent premiums are waived.
- Hospital & Surgical Benefit (H&S) rider provides reimbursement for medical expenses in case of hospitalization in a licensed hospital as an in-patient (IPD)
- Accident Death Benefit (ADB) is a one-time payment after the insured is diagnosed with a terminal illness that will considerably shorten their lifespan.

Each insured hold the different kind of riders within one policy which are under only one basic product. That means it has a different kind of riders that are often sold together. It is necessary to search the product group's pattern that suits each customer to solve this problem. Data mining techniques, such as association rules, can help extract the hidden product concurrence group from previous history purchasing of customer data. The target results are finding all the association patterns from various historical customer data features that present the customer's behavior when they bought products as a subset of frequent item sets. Most of the time, they also bought the remaining items in the same frequent item set.

Many researchers used association rules mining techniques such as the Apriori algorithm and market basket analysis, which are techniques for generating rules from a large data set based on the co-occurrences among the items. This analysis aims to learn life insurance item sets of basic and rider product concurrence to understand what customer is demanding, and these patterns can be used for recommending product sets. The concept of association rules is below.

Association rules mining results in the implication of the form X => Y (If X then Y), where X and Y are the item sets

Step 1: finding the candidate item sets which pass the minimum support[1]; Support (X => Y) = P(X and Y)
Step 2: generating rules from item sets which pass the minimum confidence[2] (accuracy)
Confidence (X=>Y) = P(Y|X) = P(X and Y)/P(X) = support (X => Y)/support(X)

Item ID	Item sets			
	Basic	Rider		
T1	Term	WP	H&S1	H&S2
T2	UL	WP	ADB	
T3	WL	WP	H&S2	
T4	UL	WP	H&S1	ADB
T5	WL	WP	H&S1	H&S2

Table 2: Sample product occurrences set

Remark: UL – universal life, WL – whole life, WP = waiver of premium, H&S1/2 = hospital & surgical benefit type 1/2, ADB = accident death benefit

[1] Support factor is an indicator that informs how much occurrence of the events X and Y.
[2] Confidence factor is an indicator that informs how many percentage of Y happen when X happen.

Table 2 is a sample product occurrence data set consisting of basic and rider products in these insurance policies. We present this example from the association rule generation with 3-itemsets rules under condition minimum support = 30% and minimum confidence = 80%. The results found that when people bought one basic product, 100% bought one waiver of premium (WP) and at least one type of hospital & surgical benefit (H&S). In terms of rider occurrences, 100% when bought accident death benefits (ADB), people bought waiver of premium (WP) and hospital & surgical benefit (H&S).

Association rules mining plays an important role in finding out the hidden pattern from basic product sell concurrence with rider product. Many researchers also use a statistical concept to analyze this point, such as K-nearest neighbor, Pearson correlation. A.B. Devale and Dr. R.V. Kulkarni discussed how the data mining method performs sophisticated classification and correlation between policy designing and policy selection [1]. They presented association rules for determined the frequent itemsets based on predefined support like riders that are often sold together with basic products. They used market basket analysis to determine the association between different policies that are sold for different purposes. In a part of classification algorithms, A.B. Devale et al. presented the classification method of K-nearest neighbors. It was used for classifying product selection in each age range and occupation, providing big pictures of how to benefit from K-nearest neighbors by doing segmentation based on income, paid premium, premium mode, and net of assurance price. They showed many examples to classify the product type in the relationship of each feature above, such as “Life security,” “Tax benefit,” “Investment” [1].

Nan-Chen Hsieh & Kuo-Chung Chu used K-means to classify customers into five clusters. It then applied the apriori algorithm as association analysis for each product group in these customer clusters [55]. This experiment investigated which function is proper to the customer's needs in a product of life insurance by extracting from knowledge rules of customer behavior and demand chain. Generating new product needs customer data and risk rate, ridership, and ridership was modeled based on limited data. The primary purpose is to share the risk and reduce premiums; after adding expense and profit, a final premium is determined with more accuracy.

To improve product recommendations' performance, we can utilize the benefit by integrating knowledge from the customer model to the product model to improve this process. Knowledge Discovery in Databases (KDD) process in data mining can merge customer data from different platforms. A.B. Devale and Dr. R.V.Kulkarni presented the statistical principles that can be enhanced the capabilities of the data mining method to improve directly into insurance procedures [1]. For example, the Pearson correlation of statistical methods to a specific correlation coefficient in positive and negative correlation was used to determine a correlation between policy designing and policy selection.

2.5.3 Risk Assesment

Risk assessment is one of the core functions of the insurance business. It estimates the likelihood and impact of risk to represent the different levels of impact with this risk. Insurance industries estimate and set the risk rating to predict the loss from accidents, health claims, or disaster rates to support the underwriting process. It set up the risk rating of the customer by placing the customer in a predetermined risk class with ambiguous names such as “super preferred,” “preferred,” “standard,” and “substandard.” This rating is estimated from customer health and physical characteristics (such as age, gender, height, weight, alcohol, drug, tobacco use, or extreme sports). This rating will directly impact the premium rate that a customer must pay to make a contract with an insurance company. The traditional method made use of actuarial statistics within the difficulty to predict base on stochastic. It is necessary to construct a risk prediction model from some techniques. Data mining and various fields of machine learning, statistics, and optimization techniques are utilized.

In this section, we give some examples of previous works from various kinds of techniques. The specific analytical techniques from data mining, such as market basket analysis, clustering technique, and statistical method, can improve the fraud claim process. IBM T.J.Watson's research center [14, 19] also developed a method called the "ProbE" (probabilistic estimation) predictive-modeling algorithm. This method uses C++ kernel rule-based models of insurance risk, where each rule represents a risk group. The purpose of this method is an analysis the competence of profitability using property and casualty insurance policy for underwriting unit. Then use it to create the predictive model for insurance risk. In the real work of actuary, the measurement of statistical confidence use to be an estimated risk parameter that can deviate from actual values. In this experiment, there are concerning points about avoiding over-fitting. It occurs when the best model of training data set perform worse when generated with new data set.

Jianbing Xiahou and Yang Mu presented the comparable result between traditional approach and data mining approach [37]. Traditional method is statistical using generalized linear models (GLMs*)*. It is an extension of general linear function $Y = X^T\beta + \varepsilon$ which consists of three factors: random factor means Y follows a normal distribution, $\mu = EY$; system factor, namely $\eta = X^T\beta$ (X^T is transport matrix of X); connection factor, namely h=m. In practice, Y does not follow a normal distribution. Researchers also comment on the general statistic method in this experiment. This task needs to process data several times to fit in practice since the section of risk variables and classification of variables are not very exact. That means the cost and complexity are huge because the input matrix is huge.

Moreover, the model obtained from the fitting may not converge. They also compared the experiment results to select the risk in classification using a decision tree algorithm to rate making into risk analysis. They expected decision tree classifier could solve the multitudinous problem of information classification. However, this experiment's results still did not satisfy since they found risk variables that make an important impact on rate making are too few in their models. In addition, the risk factor provided by the insurance agent was not reasonably rated. Yu Yan and Haiying Xie explained the usefulness of data mining in dividing credit customers into several grades, predicting customer risk, making credit ratings, and avoiding service risk [98]. They also explained in part of classification and prediction to construct a forecasting model of the future tendency of data using a decision tree, neural network, and Bayesian network.

2.5.4 Data mining challenging for knowledge extraction in insurance business

Knowledge extraction efficiency depends on the quality of data that needs all preprocessing KDD processes, especially cleansing data to handle missing data, reducing noise, and ac-counting for time series. However, we still face some obstacles of high volume and complexity of data. Many researchers are now interested in solving a problem with data mining hybrid techniques to get high-performance results.

- **High volume and high complexity in the different kinds of data**

There are different kinds of data types used in the data mining process, such as structural and unstructured data types. Structured data is easily searchable with simple, straightforward search engine algorithms, while unstructured data is the opposite. All data types are collected in different repositories such as relational databases, time-series databases, data streams, or multimedia databases. We need to integrate all data platforms into the same platform before extract knowledge by data mining technique. Most of the historical data in a company is structured data; however, the data has an enormous data stream volume. The unstructured data also plays essential roles now because some data in the company is a narrative platform. There are plenty of free texts, such as customer comments retrieved from various websites. The unstructured data needs some necessary steps to transform unstructured data to be structured data before performing data mining by some techniques such as natural language processing (NLP) and human language technology (HLT).

Pennock et al. performed a semantic classification of reviewing messages in unstructured data using N-gram's natural language processing. Then be continued data mining processes to compare the results between support vector machine (SVM) and Naïve Bayes from C|Net text corpus [22]. Hu & Liu extracted the co-occurring sets of terms from unstructured data. They build feature from these series of events and use linking rule mining to find all common sets of items. In this context, a list is a set of words or phrases that occur together. They used the apriori algorithm to find all frequent itemsets by satisfying user-specified minimum support; however, they did not construct any sequent patterns. In sum, the process of unstructured data utilization needs more complicated methods such as NLP, IR to preparing data before extract knowledge from data mining algorithms [11, 31, 32].

- **Hybrid techniques in data mining research**

To make the data mining method more powerful, many researchers presented novel methods by consolidated data mining algorithms mixed or consolidated with other knowledge such as artificial intelligence, information retrieval (IR), and NLP. For example, Nan-Chen Hieh & Kuo-Chung Chu presented the topic of enhancing customer behavior analysis by data mining techniques [55]. This research paper showed a two-stage framework of consumer behavior analysis. The key feature is a cascade involving a self-organizing map (SOM), a neural network, and a decision tree inducer. This hybrid framework used clustering techniques to preprocess input samples into homogeneous clusters and decision tree techniques to build a customer profile. This paper did the clustering customer behavior from general demographic information such as re-payment cycle, number of purchases, card age month, block code, sex, and credit line. They gave some suggestions about the clustering algorithm that affects the performance of the clustering result. The quality of input samples leads to misclassification.

Shu-Hsien Liao et al. investigated which function best suited the needs of consumers in life insurance products, drawing on the form of knowledge and specific rules from consumers and the demand chain [72]. This research used the Apriori clustering algorithm to illustrate the marketing segmentation and demand chain on life insurance in Taiwan. They also did a segment the potential of insurance buyers into similar groups using K-means and association rules based on their previous purchase.

2.5.5 Conclusion of data mining techniques in life insurance

We investigated the knowledge extraction from data mining techniques used for life insurance. The primary purpose is to understand the current customer behaviors and utilize them to recommend a suitable policy to new customers to gain their highest benefits and satisfactions. The company also has efficient customer relationship management and gains a higher persistence rate. Various data mining methods were used for knowledge extraction. Each method has different strengthen points to utilize. In this paper, we gave some examples of data mining methods to give some ideas to utilize them.

Furthermore, we found that Furthermore, we found that many researchers are interested in hybrid techniques to make knowledge construction more efficient. The correctness of knowledge extraction from data mining needs to overcome the obstacle from high volume and high complexity from the different data sources, such as structured and unstructured data generated from the diverse medium.

2.6 Life Insurance and Sentiment Extraction

2.6.1 Sentiment Analysis

Towards this, sentiment extraction is one of the main processes in the social media utilization of the social CRM concept. Theory and previous work of sentiment analysis have been reviewed. In recent years, sentimental analysis or opinion mining has been widely used to understand how people feel about a product, a service, a person, or an object of interest. Sentiment analysis or opinion mining is also known as sentiment extraction and affective rating (Binali et al., 2009). Sentiment analysis is defined as determining subjectivity in sentence and orientation in terms of negative, positive, or neutral based on the polarity strength (Binali et al., 2009).

In recent years as a sentiment analysis information extraction application, it has been widely applied to the voice of customer content such as reviews and survey responses, online and social media. It has been shown to be difficult for human readers to find relevant resources, extract relevant sentences with comments, read summaries and organize them into usable formats. This limitation triggers research on the automatic sentimental analysis or opinion mining to understand how people feel about a product, a service, a person, or an object of interest by extracting opinions or human sentiments from texts [11, 32].

Nowadays, many researchers are focusing on sentiment analysis widely. The general architecture process starts with identifying a categorization of opinion sentences into positive or negative sense and summarizing the results. The knowledge is then extracted from the assessment text consisting of different validation format that may require different techniques to perform characterization. Attributes of a subject element are tried to understand whether it is positive, negative, or neutral. Pang and Lee [2004] describe that opinion mining's primary function is to extract the opinion or human sentiment based on texts from a document analysis process. The text of the world can be divided into two main categories: facts and opinions. Facts are objective expressions about entities, events, and properties. Opinion is often a subjective expression that describes people's feelings, assessments, or feelings towards entities, events, and properties. [Bing Liu, 2010] [48]. This thesis focuses on opinion expressions that express positive or negative sentiments under CRM features.

In many cases, comments are hidden in long forum and blog posts. It is difficult for a human reader to find relevant resources, extract relevant sentences with comments, read summaries, and organize them into a usable format. Therefore, it is imperative to have an automatic search and summary system. Sentiment analysis, also known as opinion mining, has grown from this demand. It is a natural language processing issue or challenging text mining. (Bing Liu 2010)[48]. In general, sentiment analysis was reviewed mainly at three levels: Document-level: The task was to classify whether all comment papers expressed positive or negative feelings. (Pang, Lee, and Vaithyanathan, 2002; Turney, 2002). Sentence Level: The work fits into a sentence and determines whether each sentence expresses a negative, positive, or neutral feeling. Usually, neutrality means no opinion. This level of analysis is closely related to the subjective classification. (Wiebe, Bruce, and O'Hara, 1999) [7, 83], which distinguishes sentences (called objective sentences) that express factual information from sentences (called subjective sentences) that express subjective views and opinions. Aspect Level: Both the document-level and sentence-level analysis did not discover what people like and dislike. Aspect level, do a detailed analysis. Previously, aspect level was called feature-based opinion mining (Hu and Liu, 2004) [31,32]. Instead of concern with language construction (clauses or phrases, sentences, paragraphs, documents), aspect level directly points at its opinion. It is based on the idea that an opinion consists of a sentiment (negative or positive) and a target (of opinion).

In sentimental analysis and opinion mining, four major steps are

(1) Opinion identification, In the step of opinion identification, since opinions in many cases are hidden in long posts and blogs, the model needs to identify whether a sentence or a part of a text is a fact, opinion, question/answer, advertisement, or other types. Typically, two main types of information in texts are (1) facts (objective expressions) related to entities (objects, events, and their properties) and (2) opinions (subjective expressions) describing one’s sentiments or feelings on entities; even other information types include questions, answers, advertising and so forth [10].

(2) Polarity recognition, once a sentence was recognized as an opinion, we need to determine whether the opinion is positive, negative, or neutral for polarity recognition in the second step.

(3) Target identification, as the third step, is known as aspect-based or feature-based sentiment analysis [16] target of the opinion needs to be identified. The target can be an object, an individual, an organization, an event, a topic, a product, a service, or/and its associated component, attribute, and feature.

(4) Aspect or feature-based summary has to be generated from sentiments or opinions in multiple sources. So far, several researchers have studied the problem of generating feature-based (aspect-based) summaries from customer reviews of products called FBS (Feature-Based Summarization), such as those in [10, 16, 69].

Opinion identification determines the semantic orientations (positive, negative, or neutral) of opinions expressed in reviews' product features. There are many applications to support this problem such as opinion mining, summarization, and searching including the technique to utilize a list of opinion words (also called opinion lexicon) for the purpose (Bing Liu and Xiaowen Ding, 2008) [76]. They offer an effective way to pinpoint the meaning of opinions expressed by reviewers on product features. They can deal with the semantic orientations independent context by a holistic approach and aggregate multiple opinion words in the same sentence. There are many researchers study the generating problem of feature-based summaries from the product reviews called FBS (Feature-Based Summarization), which has also been implemented (Minqing Hu and Bing Liu, 2004) [48].

Previous works of Riloff and Wiebe have proposed different approaches for automatically constructing the lexicons for the feature-based opinion mining [93]. Most approaches applied some machine learning algorithms for learning the rules from the corpus. In machine learning, it has often been practical to use labeled and unlabeled examples in tandem, e.g., Nigam et al. (2000). Turney's model introduces further consideration of incorporating human-provided knowledge about language. This paper was built models that utilize all three sources: labeled documents, unlabeled documents, and human-provided information. The basic concept behind Turney's model is quite simple. The "sentiment orientation" (Hatzivassiloglou and McKeown, 1997) of a pair of words is taken to be known. Generally, many papers presented the experiment in English language corpus. Some papers presented the experiments in another language corpus, such as to propose an unsupervised lexicon building method for the detection of polar clauses, which convey positive or negative aspects in a specific domain (Hiroshi Kanayama, Tetsuya Nasukawa, 2006) [97]. Their lexicon can be expanded automatically using un-annotated corpora, and tuning the threshold values is not required. They used Japanese corpora from discussion boards. Some approaches employed morphological structures of words to extract opinion words in the Chinese language instead of the bag-of characters approach. (Ku et al., 2009).

2.6.2 Sentiment Analysis in Thai Life Insurance

While insurance companies collect large volumes of text daily through their agents, customer service centers, emails, social networks, online web communication, and so on, most cannot handle, classify, interpret or extract the essential information from such materials efficiently. As precious sources for customer analysis, such information includes policies, claims, and complaints, results of surveys, expert and health reports, relevant interactions between customers and non-customers (prospect) in social networks, etc. The insurance industry is among the ones that most can benefit from the application of technologies for the intelligent analysis of a free text (known as Text Analytics, Text Mining, or Natural Language Processing). Insurance companies have to cope with the challenge of combining the results of the analysis of these textual contents with structured data (stored in conventional databases) to improve decision-making. In this sense, industry analysts consider essential the use of multiple technologies based on Artificial Intelligence (intelligent systems), Machine Learning (data mining), and Natural Language Processing (both statistical and symbolic or semantic). The customer experience is a key element in the commercial success in the insurance sector, where differentiation between the insurers' products is not easy. Companies are trying to know their customers and their opinions by employing satisfaction surveys and directly through social networks. Text analytics permits to classify interactions according to the products or services offered, the marketing channels used, the operations employed, etc.

Besides, automatic opinion and sentiment analysis techniques enable identifying the polarity (positive, negative, or neutral sentiment) about issues or specific aspects of a product, channel, or procedure. When this type of analysis is applied to comments in open social networks, it is also possible to detect trends in the sector and identify brand perception (which concepts, activities, or entities are associated. And how we differentiate ourselves from competitors), qualify our company's corporate reputation or brand, or provide early warning of potential reputational crises. The activities that have to do with market research and competition analysis are essential to determining its product strategy [62].

To make high benefits in the real world by present the sentiment analysis results related to business topics and measure the subjectivity for giving a score on each sentiment term to show the severity of each topic detail. A company can understand the customers' minds and gain insights into its strengths and weaknesses. Customer satisfaction is one critical success of a business, so understanding customer minds in each business function are needed. Our previous research declared the group of analysis tasks as the essential functions discussed widely in the criticized website like pantip.com into three main analysis modules. First is intention analysis, second is topic analysis, and the last one is sentiment analysis.

1. Intention analysis consists of two categories

- Questions intention: we consider this category to express customers' real needs that are faced with insurance problems or asking for information related to products and services.

For example, *"ต้องการทำประกันชีวิต แบบสะสมทรัพย์ให้ลูก ขอคำแนะนำด้วยค่ะ," "I would like to make a new endowment insurance contract, please recommend."*

- Informative intention: this category refers to messages in which users provide knowledge or useful information to society. For example, "*ประกันชีวิตนำไปลดหย่อนภาษีรายได้ ส่วนบุคคลได้ หนึ่งแสนบาทต่อปี," "Life insurance premium can use for personal income tax refund 100,000 baths per year"*.

2. Topic analysis: This analysis's main purpose is to show demands or awareness in these topics for businesses to solve problems in time. It consists of three categories.

- Life insurance products: There are many kinds of life insurance products. All of them have many conditions for different purposes. The example of a basic product in different purposes such as

An endowment is periodic insurance providing for the payment date beyond the maturity date of the policy.
Whole life insurance is permanent insurance and provides lifetime coverage.
Term life insurance is for a limited time, such as ten years etc.
Universal life insurance is flexible and allows increasing premiums over time.

- Regulations: This category refers to messages or posts related to rules and conditions under life insurance concerns. All of the issues affect business units in each operation, such as the set of regulations about life insurance policy issue & rejection, the regulatory rules of claim procedure, etc.
- Service Centers: High customer satisfaction is gained from high-quality service from service centers. So this category consists of various criticized contents from a customer related to selling agents, call center, or other service centers.

3. Sentiment analysis: This intention shows the sense in customers' minds towards customers' intention related to each analysis topic. It has a positive and negative sense.

2.6.3 Hierarchical Aspect-based Sentiment Analysis

In previous work of aspect-based or feature-based sentiment analysis, there is an attempt to find out the targets on which opinions have been expressed in a sentence. Some researchers studied the summarization of generating feature-based of customer products-reviews called FBS (Feature-Based Summarization), which has also been applied [48]. In this research, we also study the CRM conceptualization to identify aspects and sub-aspects to support the business structures/process.

An important task is how to set up the granularity of sentiment level; the polarity scoring is identified. In the real situation, many words' polarities depend on the aspect (Li, Huang, and Zhu 2010), and the differentiation of prior and contextual polarity is crucial [42]. For example, 'Slow', 'Fast' words are showed in a different aspect. They may lead to be different sense and also impact the polarity of that word. Not only 'adjective' or 'verb' can declare the polarity of sentiment, but 'noun'/'pronoun' or phrase can represent the sentiment word's polarity. So in our text mining task, we utilize the Thai standard dictionary and construct a special dictionary to track the semantic of words or phrases. However, one of the most challenging issues is how to handle sentiments in several aspects in a concise, systematic, and comprehensive way.

In recent years, there is suddenly increase the interest in aspect-based sentiment analysis [42]. However, less paper is concerned about the structure of the aspects. Significantly, the aspects identifying under business knowledge need the layer of a business function to clarify each object's sentiment clearly because the components of aspect also have a relationship with sub-aspect components. For example, social CRM conceptualize components have main functions: People, Process, and Technology. What are aspects or sub-aspects under People, Process, and Technology should be declared?

From the viewpoint of technology, the hierarchical such tree structure can support sentiment extraction. Well-identified sentiment words contribute much to the accuracy of sentiment analysis. The multidimensional concept and its extension to cover online analytical processing are interesting in representing information in an abstract and implementation-independent way. There is some research related to multidimensional sentiment analysis, such as Suin-Kim et al. purpose a hierarchical aspect sentiment model (HASM) to discover a hierarchical structure of aspect-based sentiment from an unlabeled online review. They use a tree structure to represent to help the user quickly understand the significant opinions from massive online reviews [42]. Steven C. Harris et al. presents the development and initial results of a unique multidimensional sentiment analysis agent for an online learning environment to provide overall student feedback. However, this area's research still goes on because it is not easy to jointly model aspect hierarchical structure and the aspect-sentiment topics from an unlabeled corpus. [42]

Our research declares an aspect-based sentiment lexicon using aspects in the Thai life insurance industry similar to the above details and supplementing more details in the life insurance business process related to CRM functions. Then merge the sentiment analysis concept into our new approach using a multidimensional sentiment cube model to manage and visualize relations among factors (aspects/sub-aspects in a hierarchical form). A set of suitable dimensions is designed to consider standard CRM factors for the Thai life insurance industry on aspect-based sentiment analysis.

2.6.4 Sentiment Analysis Scoring

Taboada's previous task presents an approach to extract the most relevant sentences and calculate semantic orientation weighting using the SO-CAL (Semantic Orientation CALculator) concept. This task also followed Turney's basic approach to calculate orientation experimenting with the web-based search context [50].

2.7 Sentiment Extraction with Unstructure data

2.7.1 Text Mining and Natural Language Processing (NLP)

Text or message analysis that the computer is trying to understand the nature of language applied knowledge in many fields such as knowledge discovery or data mining and natural language processing (NLP), etc. (Andreas, Gerhard, Fraunhofer and Sankt, 2005). The process of knowledge extraction from large textual information to gain useful textual information using comprehensive information such as text data sets or natural languages is called text mining, where knowledge is acquired. In terms of latent data search, it is used in the analysis [41].
The characteristics of the data for analyzing latency separated into two types are:

(1) Structured data: It is called data mining analysis by data that have patterns.

(2) Nonstructural or unstructured data: Unstructured or implicit structured data, most often in the form of text or natural language, is called text mining analysis. Or the knowledge discovery process from Text (KDT).

For the analysis of the message, there is a process different from text mining. It is a structured or unstructured structure. Therefore, the process of message analysis is necessary to convert the unstructured message to be structured. Ronen and James, 2007) are called the data preparation process. The data preparation process is a necessary process that is time-consuming to obtain the appropriate structured data. This point is also different from the data mining algorithms for analysis purposes. A preparation process involves Natural Language Processing Techniques, which helps prepare the message for the computer processing for more understanding of words and sentences in natural languages. The result of the searching of latency information is knowledge or useful information.

Therefore data mining will be applied with the natural language processing (NLP) and text mining theory to improve algorithms. Data mining process consists of main processes as below:

1. Data selection
2. Data cleansing
3. Data constructing (data population)
4. Data Integration
5. Data Transformation

For the text mining task in this research was processed with NLP task, we classified the data preprocessing into three processes.

1. Feature selection
2. Text cleansing
3. Text representation

Techniques for explaining and predict data model consists of two main techniques.

1. Supervised learning such as text classification or text categorization
2. Unsupervised learning, such as clustering technique.

Model evaluation from data mining required four processing theories, which compose of 1) Text representation, 2) Natural language processing, 3) text classification and text classification evaluation as details below.

2.7.2 Text Representation

Unstructured data analysis needs to be transformed into a structured format for computer processing. The text representation process makes a substitute for a message in the form of a Vector Space Model (VSM) that is a method to represent a message in terms of feature vector space. It can choose text attributes to represent in several ways, such as words, phrases, or parts of speech (POS). Usually, text representation analyzes based on words called (bag-of-word) substitute for every word in a document with a vector consisting of a member of the value representing the word within the message. The method to represent the value representing the text message is to calculate the value for defining the message. (Atorn Numtiyagul, 2006).

- Word replacement with a occurred value or not occurred value (Binary weighting)
  - *Binary weighting =* $\begin{cases} 1 & \text{, otherwise} \\ 0 & \text{, for term represent in the document} \end{cases}$
- Term frequency (TF)
- Term Frequency – Inverse Document Frequency: TF-IDF
  - TF-IDF is the most popular way to represent text. This is because it is a simple and high efficient way of calculation result.

Term Frequency – Inverse Document Frequency (TF-IDF)

Salton and Buckley (1988) propose the method for text representation with *tf-idf* which is word frequency calculation with weight of word (w_d) calculate from *tf* multiply with *idf* value

$$w_d = f_{w,d} * log\left(\frac{|D|}{f_{w,d}} \right)$$

when $f_{w,d}$ is tf (term frequency) which is frequency value of word (w) in document (d) multiply with logarithmic scale of number of document (D) which divide by the number of documents are found words (w). It means the inverse document frequency : idf (Juan, 1999).

However, this method has a limitation of evaluation performance for document classification. Because some words inside documents are not a good representative of a document, “stop word removal” and “word stemming” are methods for classified word qualification; these methods can increase text classification efficiency. Moreover, it can reduce document volume down 30-50%.

- Stop word removal is the word that has a high opportunity to occur in the document and does not useful for text classification such as conjunction, preposition, ending word, etc.
- Word stemming is the word that has the same root word; however, the pattern is changed, such as "eat, ate, eaten". They have the same root word, "eat". However, the Thai language does not have that word characteristics. Nevertheless, the Thai language has the same word pattern in different meaning

or different word patterns in the same meaning, such as 'กิน/eat' and 'รับประทาน/eat'.

Vector space model that uses single or phrase of words to represent a word in documents has many previous works proposed the techniques about feature selection by considering the word occurrence such as information gain (IG), mutual information (MI), chi-square, etc. Minqing Hu and Bing Liu (2004b) proposed the method that uses association rules mining for analysis of the frequency of the relation among features start from two words co-occurrence onwards [31].

2.7.3 Association Rules Mining

Association rules mining is a rule to finding out the relationship of data more than of equal two sets within the same big group of data. This technique was proposed by Rakesh Agrawal and Ramakrishnan Srikant (1994). One of the useful techniques of association rule mining is the 'market basket.' The process that constructs the data relationship rules separate into two processes: 1) Finding the frequency item set and measuring the relationship among features with 'support' values. If the association of some data sets has a support value more than the 'minimum support' value called 'Frequency pattern' or 'Frequency item set.' The next process is 'Association rules' identification, which is a process using the association under minimum support value to construct the rule. If some of the association of dataset has 'confidence' value more than 'minimum confidence' which is limited called 'Association rule.'

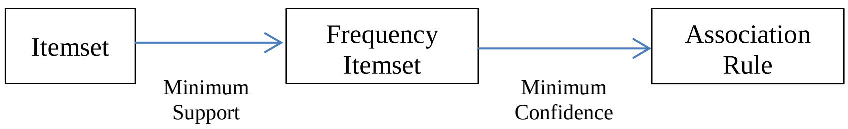


Figure 3: The process of association rules mining

Support is an indication of how frequently the item set appears in the dataset. (Kenneth and Narciso, 2007)

$$Support\ (AB) = \frac{transactions\_contain\_AB}{total\_transactions}$$

Confidence is an indication of how often the rule has been found to be true (Kenneth et al., 2007)

$$Confidence\ A \rightarrow B = \frac{Support(AB)}{Support(A)}$$

2.7.4 Natural Language Processing (NLP)

Natural language processing (NLP) is one of the most important information technologies which try to understanding complex language utterances is also a crucial part of artificial intelligence. Human language is used to communicate with a computer in a structured language that can be processed immediately, such as PHP, Java, or C ++. However, human language has its structure called natural language, which is an unstructured structure.

Natural language processing is divided into the following analysis levels: (Christopher and Hinrich, 1999)

1. Morphological analysis is a word analysis level that makes use of Lexicon knowledge.
2. Syntactic analysis is Part-of-Speech of words analysis level, which is the basic information for examining language's grammatical rules.
3. Semantic analysis is an analysis of the meaning of words in a sentence.
4. Discourse integration is considered the meaning of a sentence with an adjacent sentence.
5. Pragmatic analysis is the interpretation of Sentences in which the speaker wants to convey.

The language analysis process starts at the lowest level, morphological analysis (word level) until pragmatic analysis can explain the language structure elements. That is word, morpheme, phrase, noun phrase, verb phrase, sentence, and grammar. An analysis in higher-level more than pragmatic requires a knowledge base of a word such as WordNet. The Thai language is frequency analysis in morphological analysis and semantic analysis because Thai WordNet is underdeveloped and has some limitations (Alisa et al.m 2010) [17]. Morphological analysis and semantic analysis in Thai language analysis have more complicated than in English. The Thai language does not have any stop wording in a sentence, including writing style is to continue until becoming one passage. Therefore the process of managing the Thai language before analysis requires three necessary steps: 1) Tokenization, 2) Part-of-Speech tagging, and 3) Syntactic analysis.

2.7.5 Word Segmentation

Research about word segmentation is separated into three techniques. There is 1) a Rule-based approach, which is a method considering character or alphabet. This method is the most convenient and fast; however, it cannot solve ambiguous words problem. 2) Dictionary approach: This method makes a segment of a word using 'longest string matching' with 'maximal matching.' This method is more accurate than the rule-based; however, it still cannot solve all of the ambiguous word problems. 3.) Corpus-based approach: this approach use statistic to calculate the probability of word occurrence. This method gets the most accuracy; however, there is a limitation based on the accuracy of corpus database.

2.7.6 Part-of-Speech Tagging (POS tagging)

Part-of-Speech tagging in the level of syntactic makes a better understanding of the sentence meaning. This process makes specific the function in each word that present in the grammatical pattern. The word's functions consist of N-noun, PRON-pronoun, V-verb, AUX-auxiliary verb, ADJ-adjective, PREP-preposition CONJ-conjunction, DET-determiner, CLAS-word class, NEG-negative/reject the term, END-stop word.

2.7.7 Grammatical Syntax Analysis

Syntax or Parsing is the process that explains the sentence structure with Grammar formalism to define the word pattern in a sentence. It describes in a Tree diagram and Labeled bracketing. One of the popular gramma formulas is ‘Context-Free Grammar: CFG). Context-Free Grammar is one of phrase structure grammars. PS which does not concern the word meaning but consider in the ordering of word function from left to right and split step with two symbol types 1) Non-terminal symbols such as Part-of-Speech (POS) or chunks 2) Terminal symbols refer to a word in a dictionary. Generally, the Thai language has a lot of ambiguous meaning and conversational language in sentences. The challenge problems in the Thai language are managing the long sentence and separating wording with the correct definition. The Thai language is the one Asian language has difficulty in clarifying the end of each sentence. Many papers about Thai language classification present in terms of the word segmentation by using the dictionary base. This method may face with the ambiguity of the word. So another technique to do word classification is using statistics base on the various corpus.

Asanee et al. proposed an approach for Thai segmentation by using decision tree learning [4, 5]. The next research has the experiment one of closely related to this paper is Constructing Thai Opinion Mining Resource: A Case Study on Hotel Reviews of Haruechaiyasak et al. suggest a Thai language resource of constructing framework for feature-based opinion mining that extracts lexicons, including domain-dependent and domain-independent, based on syntactic pattern analysis. This experiment found that polar words could be extracted more easily than sub-features [36]. It may cause sub-feature extraction, and polar words are not straightforward meaning owing to it is quite difficult to deduce from sentences of the corpus. The other task of Haruechaiyasak et al. proposes an S-Sense framework for analyzing sentiment on Thai social media on intention, sentiment, and language usage in Thai texts [17].

Benea et al. (2008) is automatically generated resources for subjectivity analysis in a new language using machine translation and standard Naive Bayes and SVM leveraging on the resources available in English [88]. This experiment exemplifies the technique of Romanian and Spanish. In terms of negation in Sentiment Analysis, some papers presented in the survey pattern find out how to solve the ambiguous meaning problem in negative polar words (Michael Wiegand et al., 2010) [51]. This research also used an approach to exploit linguistic knowledge to improve topic and sentiment classification in online opinions related to Thailand's life insurance business. This approach automatically assigns appropriate labels on a proper class by training the classification models and language resources with linguistic clues to extract sentiment from texts in section 3.

2.8 Online Analytical Processing (OLAP) & Multidimensional Sentiment Cube

2.8.1 Online Analytical Processing (OLAP) and Multidimensional Cube

Online Analytical Processing (OLAP) is one of the database technologies, which has previously attracted many researchers' interest. OLAP is based on a multidimensional data view supported by a multidimensional database (MOLAP) or a relational engine (ROLAP). The main characteristics of OLAP applications are (a) data on multidimensional view and (b) Data analysis through queries to guide the data (Panos Vassiliadis, 1999).

The multidimensional data view recognizes which information is contained in a multidimensional array (or Hypercube, cube). A cube is a collection of data cells organized by dimensions. Dimensions are defined as the features structure. Each dimension has an associated hierarchy of levels of user's analysis viewpoint. For example, the time dimension can be derived as Year, Month, Week, and Day as a hierarchy. Measurement is variables, metrics, or facts which stand for the measurements value [90]. The traditional multidimensional model is a multidimensional fact-based structure generated from complex queries [39]. A fact set up from analysis viewpoints of some concerning events called 'dimension' (e.g., service touchpoint, service process) can be quantified from measures called 'measurement' (e.g., number of complaint words). The measurements can be aggregated into different levels of concepts across different layers of multi-dimensions as hierarchical form. A data cube provides multidimensional perspective viewpoints though multi-measurements combination to multi-dimension values called 'multidimensional cell.' An analysis point of view can vary by typical OLAP operators such as 'drill-down,' 'roll-up,' 'slice and dice,' 'pivot' or 'top-k selection.'

2.8.2 Sentiment Cube

Recently, to catch the summarized figure of sentimental analysis, a sentimental cube has been introduced by Umeshwar Dayal in 2012 [84] [85]. To combine multidimensional cube with sentiment analysis theory, Umeshwar Dayal defined patent (US20120197950 A1) of sentiment cube that “A sentiment cube system is a disclosed system which stores sentiment elements inside. A sentiment cube data structure is having a set of cells arranged by a set of dimensions. The system includes a computer programmed with executable instructions that operate a set of modules. The modules comprise a sentiment storage module that receives sentiment values associated with a set of entity features and then populates a hierarchy of the cells in the sentiment cube with the sentiment values. Sentiment analysis modules affect a set of operations on the sentiment cube [85]”.

Sentiment cube emerges in the era of social media, which raise sentiment analysis techniques. The sentiment is deployed to be one of the analysis dimensions. That makes it more potent on analysis viewpoints. Start from the definition of a sentiment cube system, this study illustrates our proposed design of a sentiment cube for Thai life insurance business using the social CRM and its practice use with three main properties, i.e., multi-dimensions, hierarchical dimensions, and multi-measurements. A sentiment cube system is proposed to store sentiment elements inside, together with their target aspects. Its data structure consists of a set of cells, organized in the form of dimension structure. The system keeps sentiment values associated with a set of entity features (aspects) and then populates a hierarchy of the cells in the sentiment cube with the sentiment values. The system allows a set of operations on the sentiment cube, such as drill-down, drill-up, slice/dice, pivot, top-k selection, etc. Previous work-related to the multidimensional cube on unstructured data, many researchers started the experiment on unstructured data such as text documents, chat blogs on Twitter, etc. For example, Frank, Olivier et al. presented the OLAP multidimensional concept model without facts. This model is based on a specific dimensions concept and is used for the analysis of multidimensional documents. They also provided the operation set of cube explanations [69]. Xiong Liu et al. discussed a text cube approach to studying different kinds of human, social and cultural behavior (HSCB) embedded in the Twitter stream. They showed how to organize data from text to multiple dimensions and hierarchies and make visualization with statistical reports and perform online analytical processing [2]. Duo et al. proposed the combination of topic cube and probabilistic topic modeling by enable OLAP on the dimension of multidimensional text database. They also presented two heuristic aggregations to accelerate the iterative EM algorithm by leveraging the models learned on component data cells in a good starting point of iteration for topic models estimation.

2.9 Customer Satisfaction in Life Insurance

There are many studies related to customer satisfaction in life insurance with demography concepts as follows:

Demography Characteristics Concepts

Demographic characteristics include age, gender, family size, family status, income, occupation. These studies are criteria that are commonly used in market segmentation. Demographic characteristics are important characteristics and measurable statistics of the population that help determine the target market, including easier to measure than other variables. With different demographic characteristics having different psychological characteristics, the important demographic variables are as below

1. **Age** is a factor that makes people different about thoughts and behaviors. Younger people have liberal ideas. Adhere to ideology and look more optimistic than those who are older. Many older people have conservative ideas adhering to cautious practices, more pessimistic than younger people because of their different life experiences. The product will meet the needs of varying age groups; marketers, therefore, take advantage of their age, which is another demographic variable of the market segment.
2. **Gender** is also an important variable in market segmentation. Gender variables have changed in consumption behavior, especially women. Gender differences make people have different communication behaviors; females tend to receive more messages than males, while the male desires to create a relationship between sending that message. Also, cultural values and attitudes define the different roles of both sexes.
3. **Education** is a factor that makes people have different ideas, values, attitudes, and behaviors. Highly educated people have a significant advantage in being suitable recipients because they have a good understanding of communication and consider enough reason. In contrast, people with low education tend to find non-radio, television, and print media.
4. **Social and economic status** means occupation; the person's income and social image influence the message recipients. Each person has a different culture, experience, values, attitudes, and goals.

Attitude Concept:

Attitude refers to assessing satisfaction or dissatisfaction of a person, emotional feelings, and practical trends that affect a particular idea or a thing or referring to a person's feelings for something. Attitude is an influence on faith. At the same time, belief has an impact on attitude. The attitude is caused by the information that each person receives from the experience about the product or the person's mind, including the relationship with the reference group, such as father, mother, friend, a leading person in society, etc.

Experience Concept:

Experience means an experience that has been caused by an action or seen. Experience is valuable in learning every aspect of the experience that affects the creation of art. There are two types as follows: 1) Direct experience is an experience that we have encountered or touched by ourselves, found ourselves, acted ourselves, heard and listened to ourselves 2) Secondary experience or also known as "Indirect experience," is another experience that has been inherited or acknowledged.

Customer Satisfaction Concepts:

Customer satisfaction is one of the critical factors that will make the company successful. Making customers satisfied is essential because the advancement of service is a crucial factor. If the number of additional users increases, customer satisfaction will make the business gain more market share. The high volume of purchase repeatedly is referenced business that will lead to better profitability (Barsky, 1992). Person (1993) has given meaning to customer satisfaction; that is, the level of customer feelings toward products or services, whether the product or service can meet the needs of the customers' expectations. It gets better customer satisfaction and loyalty to products and services, causing the behavior of repeat purchases or additional services. Moreover, customer satisfaction will make the customer tell the closed person, such as family and friend.

To summarizes previous work about customer satisfaction. Many previous studies have focused on the indicators that influence customer satisfaction in life insurance. However, these studies had their objectives. For example, Kuhlemeyer and Allen (1999) explored consumer satisfaction relevant to the purchase behavior of life insurance products and compared the service provided by a broker or agent to service with no broker or agent support [43]. They attempted to identify a benchmark for customer satisfaction for life insurance products, agents, and providers. Sogunro and Abiola (2014) studied the measurement of customer satisfaction among multi-attribute products and services based on product purchase in Logos State, Nigeria. Their study recommended that life insurance service providers in Nigeria continue to conduct customer surveys on their products to identify each product's aspects that created dissatisfaction [75] (Sogunro & Abiola, 2014). Subashini and Velmurugan (2016) studied policyholder satisfaction with various factors relating to life insurance products in Coimbatore District, India. They chose this area because insurance companies were facing a significant problem of lapsed policies [78]. Some researches highlighted the importance of customer satisfaction on the service quality of life insurance; for example, Siddiqui and Sharma (2010) administered a questionnaire survey using confirmatory factor analyses by investigating six dimensions of service-quality: assurance, competence, tangibles, personalized financial planning, corporate image, and technology. They used structural equation modeling to assess the results and proposed a framework for appropriate action to satisfy customers through quality services [73]. High customer expectations in a competitive business environment led Samarasinghe et al. (2018) to focus on customer satisfaction in terms of service quality in Colombo District, Sri Lanka. Their research investigated the factors influencing the satisfaction levels of customers [71].

Kannan (2018) also studied customer satisfaction with corporate life insurance with particular reference to Chennai's city, in India. The main aim of that study was to discover customer satisfaction with life insurance companies based on primary data from a questionnaire survey of 150 policyholders. They concluded that insurers needed better understand the customer requirements for such policies and increase promotions, such as advertising. Besides, greater awareness was needed among illiterate and rural groups [40]. Nguyen et al.(2018) presented the determinants of customer satisfaction and loyalty in Vietnam's life insurance industry based on a questionnaire survey on factors that would enhance customer relations, image, quality, and added value to create a sustainable life insurance business in Vietnam [59]. (Nguyen et al., 2018). Basaula (2017) studied the awareness of customer satisfaction with the claim settlement process in Nepal using questionnaires. The results indicated that many respondents gave neutral scores on claim satisfaction. They required faster claim settlement; further, the study stated that the government should focus on creating life insurance awareness [8].

Many researchers have examined the significant indicators that affect customer satisfaction from various perspectives. Coviello and Trapani (2012) proposed a conceptual for investigating the customer retention using the satisfaction on customer relationship and quality. Their study looked at the relevant concept of client satisfaction derived from the aggregation of individual experiences in different situations, including psychological states and customer attitude toward service as the impact on satisfaction levels [21]. Bakar, Soykan, and Acar (2018) measured life insurance knowledge among students of the insurance and risk management department at Dumlupinar University in Turkey. Although the research was not related to satisfaction levels, it measured the students' basic life insurance knowledge. It found that even those who wished to work in the insurance sector had very little knowledge of basic life insurance. This indicated the need for better education and knowledge sharing with students. Further, it implied that insurance knowledge would influence the life insurance market and, therefore, more knowledge support was necessary [6].

Research on the life insurance market is also important in a developed country such as Japan, as the country experiences depopulation due to aging and a declining birthrate. Thus, life insurance companies face a shrinking market. Tomoki Inoue (2014) studied various topics relevant to customer behavior changes based on information from the internet and found that the internet will impact growth in the life insurance market. Therefore, it is important to increase our understanding of consumers' actual actions, including factors and behaviors that affect customer satisfaction, considering the increase in social media [36].

In addition, scholars have also investigated customer satisfaction in the Thai life insurance industry. For example, Phromsuwan (2011) studied customer attitude and satisfaction with buying life insurance from a telemarketing channel from a company called "Siam Commercial New York Life" in Bangkok. That study focused on customer demographics and attitudes toward the service provider [66]. Wongwiratchit (2015) also studied customer satisfaction with life insurance in Thailand. The study evaluated customer satisfaction using path analysis on customer expectations, perceived value, perceived quality, and customer loyalty [94]. Other studies on Thai life insurance have focused on satisfaction factors that impact life insurance purchases, such as Chompuphan (2014), who studied factors affecting the intention to buy long-term life insurance in Northeast Thailand [20].

In all, we found many studies on life insurance in emerging markets. Customer satisfaction has been shown to depend on customer opinions in specific areas. Thus, researchers have investigated customer satisfaction in many relevant areas, such as purchasing life insurance products, service quality, customer loyalty, claim settlements, service channels, and telemarketing. These studies have all examined influential factors and customer satisfaction levels using questions based on service providers' expertise using inside-out knowledge. Our research utilizes outside-in knowledge by tracking customer dissatisfaction expressed in messages on websites to extend a current analysis. They are then analyzing the extracted knowledge for service providers ‘benefit.

Chapter 3

Implementation

This section presents the background and idea of implementing customer sentiment knowledge management and its process in Thai life insurance. The concept is explained as a knowledge management process, including the implementation of support tools; the multidimensional sentiment cube design is described and illustrates detailed implementation, including information extraction and sentiment calculation. Besides, the sentiment knowledge sharing using a questionnaire is explained.

3.1 The Concept of Customer Sentiment Knowledge Management

Customer knowledge management is the process for managing knowledge related to the customer, such as customer profile and customer activity. In the social media era, customers prefer to express feelings on social media, both positive and negative opinions. However, a company can utilize and solve the problems in time with negative sentiment extraction. In addition, customer sentiment knowledge management is also the knowledge management that exploits the benefit from customer sentiment such as product improvement strategy based on positive opinion, complaint management system based on negative emotion.

Knowledge may arise from the transmission of experience or the analysis and synthesis of information. Knowledge creation is set up from knowledge co-creation. Knowledge on customer sentiment is the knowledge arises on the particular subject related to customer sentiment on products or service. To manage the customer sentiment knowledge has significant to business sustainability because it raises customer satisfaction.

However, extracting customer minds from messages spreading in social media is not an easy task because it requires language and analysis technology based on business knowledge. There is one concern from sentiment extraction from social media; it is entirely unclear a characteristic of a person who expresses that kind of sentiment on social media as anonymous sources.

Our methodology defines customer sentiment from the information obtained from the Outside-In and Inside-Out management approach of the organization. The Outside-In approach is guided by the customer's experience and their opinion that customer value creation, customer orientation, and customer experiences are the keys to success. The value is a consequence of listening and providing value to customers. The Inside-Out approach is guided by the practice, experience that the inner strengths and capabilities of the organization will make the organization prevail. The knowledge co-creation from both outside-in and inside-out information related to customer sentiment benefited business, especially process improvement, product and service evaluation, etc. However, the utilization of that kind of sentiment knowledge requires knowledge management to make it more powerful and practical in real practice.

The difference between Customer knowledge management and Customer sentiment knowledge management:

1. **Customer Knowledge Management**

- Disciplinary partners: customer relationship management and knowledge management approaches.
- Perspective: the interface of customer and the inside of an organization.
- Key actors: customers and employees.
- Key communication context: collaboration between customer and organization.
- Conceptual focus: What is the customer knowledge from different sources and types?
- Key processes: utilize the customer knowledge of customer and organization.
- Goal: adapt the knowledge from customers and organization to support CRM efforts.

2. **Customer sentiment knowledge management**

- Disciplinary partners: CRM, knowledge management, and sentiment analysis approach.
- Perspective: customer interface, inside an organization, and social media interface.
- Key actors: employees, customers (anonymous), and customer (specific in demographic profile).
- Key communication context: collaboration between customer-specific related to customer sentiment and the organization to increase customer satisfaction.
- Conceptual focus: What is the customer knowledge from different customer sentiment?
- Key processes: utilize the customer sentiment knowledge on the organization by requiring the transform process between customer sentiments to business knowledge using technology.
- Goal: learning from customer and utilize in the organization to get process innovation or improvement to increase customer satisfaction.

3.2 The Methodology of customer sentiment knowledge management

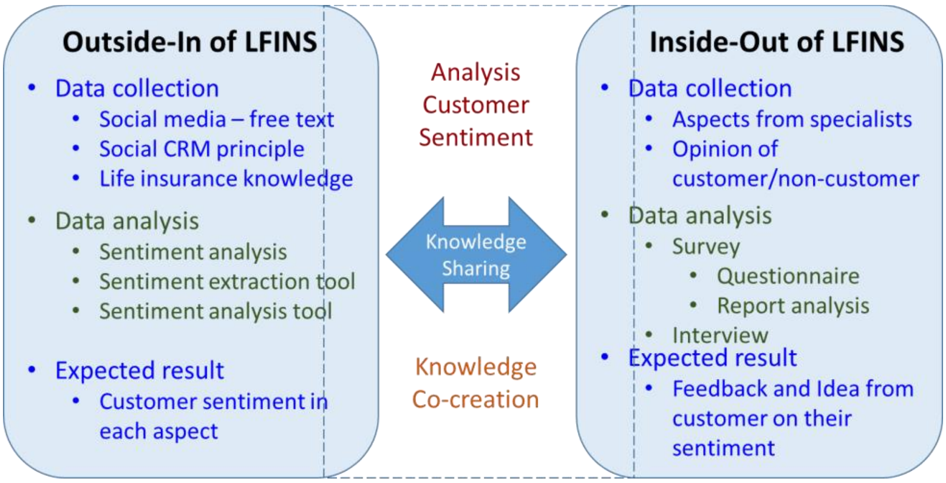


Figure 4: The methodology of customer sentiment knowledge management

Refer to the previous passage; our methodology defines customer sentiment from the information obtained from the Outside-In and Inside-Out management approach of the organization to get knowledge for finding out the methods or processes to support life insurance increase customer satisfaction.

The methodology starts from the **outside-in approach** of life insurance, which **expects** to get customer sentiment in each service aspect of life insurance from social media. **Data collection** from this process will get free-text from various critiqued websites in life insurance service merge with the social CRM principle and life insurance knowledge. In the data collection step, we will get data in text forms and require technology such as text preprocessing, text mining, natural language processing (NLP), and sentiment analysis to extract customer sentiment, (*refer to Figure 4: Step 1 – Identify and access*). Next, **Data analysis** deploys the sentiment analysis concept for analyzing customer sentiment results. In this step, we use the "***Sansarn Tagging tool***," which helps extract the seed of words (keywords) as a sentiment extraction tool. We propose a "***Multidimensional sentiment cube***" as a novel sentiment analysis tool that can provide unlimited analysis viewpoints as the concept of unlimited cross operations of online analytical processing (OLAP), (*refer to 3.3: Step 2 – Extract and validation*). After that, we will get the customer sentiment information such as sentiment score, sentiment ranking level, etc. in each analysis aspects, (*refer to 3.3: Step 3 – Utilize and analysis results*).

Next is the knowledge sharing process between outside-in and inside-out information. For the **inside-out approach**, we do the knowledge sharing process using the survey by a questionnaire and interview in the **data analysis step**. So the **data collection** consists of aspects from specialists, opinion of customers/non-customers in Thai life insurance service. We used the sentiment analysis results which we got from outside-in information (critiques from social media) to be questions in the questionnaire for the knowledge co-creation process. We utilize customer dissatisfaction issues which we got from previous steps, to be the highlighted issues for proving in real practice and finding out the appropriate problems solving methods in a questionnaire. They consist of service evaluation issues, claim issues, policy cancellation issues, and other misunderstanding issues, (*refer to 3.3: Step 4 – Sharing*). To capture knowledge sharing between customer sentiment and respondents who replied to the questionnaire, the results of customer sentiment knowledge sharing got from their experience, (*refer to 3.3: Step 5 – Capture and learning*). Knowledge co-creation occurs in this step, the expected results are feedback and idea for problem-solving methods from customers, non-customers and expertise in life insurance area, (*refer to 3.3: Step 6 – Create and leverage*). In the final step is implementation, we can find more knowledge after implement in the real practice then we learn from mistake and improve it. In this step, the efficiency of monitoring process plays important role to catch up the riddle and require the process improvement, (*refer to 3.3: Step 7 – Implementation, Step 8 – Learning and monitoring*).

3.3 Customer Sentiment Knowledge Management Process

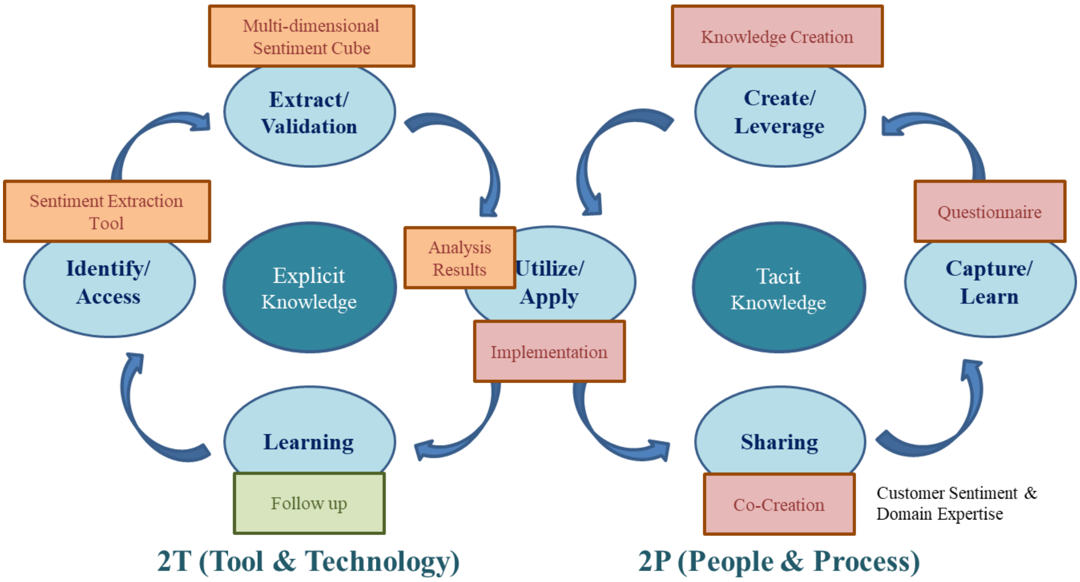


Figure 5: Customer sentiment knowledge management process

Step 1. Identify and Access

In a part of store and access, it is necessary to identify a related knowledge such as a source of sentiment. In the concept of social CRM, we have attention to find customer satisfaction using customer sentiment from social media using technology. Then this task has developed a tool named "*sentiment extraction tool*". However, develop this tool requires business understanding. In this step (Step 1) has two tasks. There is Task A: Life insurance business understanding with social CRM and Task B: Sentiment Extracted Tool - Sansarn Tagging Tool as following.

1.1 Life Insurance Business Understanding with Social CRM

Life insurance process is the one of a process which closely the relationship between the service process of company and customer who has left comments related to service on social web-blogging. To identify the life insurance service process, we have to review the service process on this business in the principle of general practice. We explained in terms of 'Customer life cycle' concerns on the management information, which consists of three stages: customer acquisition, customer engagement, and customer retention. We select only the processes which highly involve customers' sentiment on social media.

Stage1: Customer Acquisition (Pre-purchase → purchase process)

A service on customer life cycle starts from the company find out a list of prospect (expected customer) to offer new plan (product) or contract. Service touchpoint, a company representative, introduces the customer's proper plan to meet customers' requirements and their lifestyles, such as financial status, health condition, premium payment, or sum assured earning expectation. Service touchpoint and prospect should make understand and agreement on exceptions and conditions under a life insurance contract. After the prospect decided to apply for a new life insurance contract, a new application will be submitted to the underwriting department to verify customer qualification and then issue a new policy contract and deliver to the customer by service touchpoint. The acquisition stage will be a success by providing product and service prominently with innovation to attract customers' minds with freewill service [45].

Stage2: Customer Engagement (Purchase → service support process)

After new policy contract was approved, customer status is changed from prospect to be insured of a company due to, the customer will obtain service under contract conditions and life insurance law such as service on claim assessment as life insurance proceeds. Insured acquired an excellent service from service touchpoint when new procedures occurred. For example, service touchpoint should be easy to contact a customer, help to solve problems related to insurance service including general support in operation service such as change policy status (make extended term insurance (ETI), reduced paid-up insurance (RPU) or cancel policy), make loan from policy when cash value of policy occurred.

(***Note****: insured is a policy owner or person whom the insurer (Life insurance Company) promises to pay a designated beneficiary a sum of money (the benefit) in exchange for a premium, upon the death of an insured person under policy contract conditions) [50].*

Note*: insured is a policy owner or person whom the insurer (Life insurance Company) promises to pay a designated beneficiary a sum of money (the benefit) in exchange for a premium, upon the death of an insured person under policy contract conditions) [50].*

Once the insured requires to make a claim such major claim (death claim) or a minor claim (illness claim), the customer has to contact the insurer to request compensation as a life insurance contract. The company department of service touchpoint should provide professional service to the insurer by considering that claim, give advice, and decide whether the insured's policy cover the hospital costs as soon as possible. To provide experienced service and follow-up with care to the insured makes the engagement process better and increases customer loyalty [45, 83].

Stage3: Customer Retention (Re-purchase process)

Customer Retention main point is to keep insured with active status or recall ex-customer by resale with a new product or retain customer before leaving the company. For example, when policy status changed to be mature (end of the contract), the insured became inactive. However, retaining customers' needs to get satisfaction from a customer so listening to customer's voices to find out dissatisfaction points on the current business process is important.

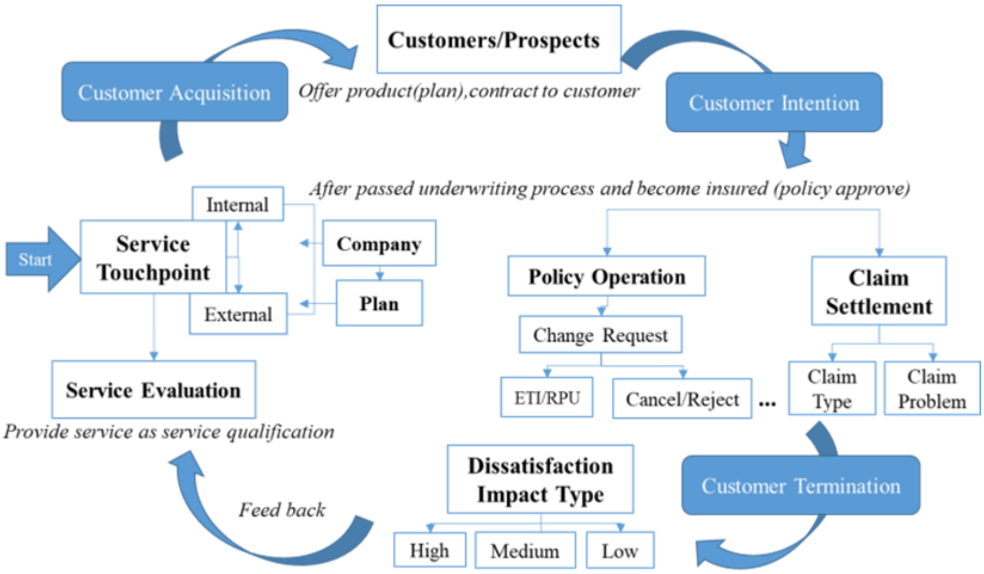


Figure 6: Life insurance process across to CRM

This demonstrated scenario is a life insurance process across to CRM function, consisting of People, Process, and Technology. An analysis of the CRM strategy consists of customer acquisition, customer intention, and customer termination. Life Insurance processes under service unit are service touchpoint, operation department including the process of new policy approval, policy change request, and claim assessment. All of the units are analyzed by various systems such as CRM system, DWH/OLAP, BI, Sentiment analysis with social media as Table 3 and Figure 6.

Three CRM strategies are concerned under customer life cycle of CRM conceptualize related to life insurance business process: [Table 3]

1. Customer acquisition Customer acquisition is a strategy to extend a business's market share by approaching new customers and making ex-customer returns. Using social media channels, it has many ways to manage, such as launch a new campaign to a direct-market group. It is also essential to understand the current situation of customer sentiment and current problems in the acquisition stage for all of them. To assign a proper aspect, it needs to study the business process. The process in life insurance of the acquisition stage as below: [Table 3 on item 1.1 – 1.5]

1.1: Start → **service touchpoint** (external such as agent or agency, internal such as call center: Refer to table.1) →*offer plan which appropriate to requirement* → *provide important information & special exception* → *consider health condition & financial performance* → **customer** [Table 3 on item 1.1, 1.2, 1.3]

1.2: Customer → *reveal fact (personal information, health condition)* → *make understand on new plan information with exceptional conditions* → *sent a new policy contract to approval process* → **service touchpoint & operation department** [Table 3 on item 1.4, 1.5]

Table 3: Life Insurance process related to CRM components and strategies

Remark: ✓means this life insurance process has relationship with CRM component, ○ means this life insurance process occasionally has relationship with CRM component, **-** means this life insurance process does not have any relationship with CRM component.

CRM Comp.	Life Insurance Process	People				Process				Technology				
		CUST	Serv. TP.		OP Dept.	New P.A.	P.C.R.	Claim Ass.		CRM	DW	BI	SA	SMA
			EX	IN				Minor	Major					
Customer Acquisition	1.1 Propose appropriate plan as customers' requirement	✓	✓	✓	-	✓	-	-	-	✓	✓	✓	✓	✓
	1.2 Explain clearly an important information & special exception	✓	✓	✓	-	✓	-	-	-	○	○	○	✓	✓
	1.3 Consider and reveal current health condition & financial performance	✓	✓	✓	-	✓	-	-	-	○	○	○	✓	✓
	1.4 Make understand on new plan information with exceptional conditions	✓	✓	✓	-	✓	-	-	-	○	○	○	✓	✓
	1.5 Submit a new policy contract to approval process	✓	✓	✓	✓	✓	-	-	-	✓	✓	✓	✓	✓
Customer Retention	2.1 Provide customer support when policy 's requirement is changed	✓	✓	✓	✓	-	✓	-	-	✓	✓	✓	✓	✓
	2.2 Support customer of claim assessment with 3Cs verification	✓	✓	✓	✓	-	-	✓	✓	✓	✓	✓	✓	✓
	2.3 Follow up problems and customers' sentiment	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	✓	✓	✓
Customer Termination	3.1 Sue on court, make petition for OIC , make an accusation to agent	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	✓	✓	✓
	3.2 Make policy cancellation and make policy reject	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	✓	✓	✓
	3.4 Feedback in complaint system and/or express sentiment on social	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	✓	✓	✓

Table 3. Three basic components in customer relationship management process, where CUST="Customer", Serv.TP = "Service Touch Point", EX = "External", IN = "Internal", OP Dept. = "Operation Department", New P.A. = "New Policy Approval", P.C.R. = "Policy Change Request", Claim Ass= "Claim Assessment", CRM = "Customer Relationship Management", DW = "Data Warehouse/OLAP", BI = "Business Intelligence", SA = "Sentimental Analysis", and SM = "Social Media Analysis"

From Table 3, there are ✓sign strongly related to People-Service touchpoint & customers on the acquisition process. Hence, an analysis viewpoint requires some indicators to monitor service touchpoint performance. This analysis task provided a "service evaluation factor" of service touchpoint for evaluating their performance. The service evaluation factors consist of reliability, consistency, dependability, responsiveness, competency, access, courtesy, communication, knowledge, credibility, security, and understanding. They are one of the main aspects of sentiment analysis in our design.

Aspects identified in the acquisition stage are ‘Service touchpoint’ and ‘Service evaluation factor,’ including sub-aspects identified under their aspects. [Table 3 on item 1.1 – 1.5, Table 4 on item 3&4]

The service touchpoint aspect has two sub-aspects. [Table 4 on item 3]

- Internal service sub-aspect [Table 4 on item 3.1] has three sub-aspects. [Table 4 on item 3.1.1-3.1.3]
- External service sub-aspect [Table 4 on item 3.2] has five sub-aspects. [Table 4 on item 4, 4.1-4.5]

The service evaluation factor aspect has five sub-aspects. [Table 4 on item 4, 4.1-4.5]

2. *Customer intention*

Customer intention is one strategy to keep a current customer by understanding customers’ minds and solving misunderstandings between customer and business. Using social media, finding dissatisfaction of company service to help to solve the problems is needed. The process in life insurance of intention stage as below:

When a new policy is approved, customer status is changed to be ‘insured,’ then policy protections and services are started under coverage. [Table 3 on item 2.1 – 2.3]

2.1: Customer → *require support from service touchpoints (make “policy change request” or make “claim assessment” process)* → *contact and request for service* → **Service touchpoint** [Table 3 on item 2.1, 2.3]

2.2: Service touchpoint & operation department → *support follow standard procedure* → *claim assessment verify 3Cs* → *return or reject payment to customer as coverage rules* → **Customer** [Table 3 on item 2.2, 2.3]

- Policy change request will occur when customer’s requirement is changed, for example, require new rider plan, request to change policy status such as make extended term insurance (ETI), reduced paid-up insurance (RPU) or cancel a policy, etc. [Table 3 on item 2.1, 2.3]

Note: *ETI is a clause under many policies that gives the option of continuing the existing insurance for a period based on the contract's cash value. RPU is a policy whose cash value can be used to buy paid-up insurance in the highest amount it can afford.* ***Source:*** *http://www.businessdictionary.com/definition*

- The claim assessment process is one of the important processes in life insurance, which is the process to request the insurer to pay to return coverage of loss under the policy's terms. Payment consideration is under "Claim Assessment Triangle Theory (3Cs)," which consists of three elements, cause of the claim, coverage terms, and contractual legality [44]. [Table 3 on item 2.2, 2.3]

Aspects identify on intention stage is ‘Process and Operation’ including sub-aspects also identify under their aspects. [Table 3 on item 2.1– 2.3, Table 4 on item 5]

- Process and Operation aspect has three sub-aspects. [Table 4 on item 5]
  - Policy Operation sub-aspect [Table 4 on item 5.1]
  - Claim Operation (type) sub-aspect [Table 4 on item 5.2] has two sub-aspects. [Table 4 on item 5.2.1-5.2.2]
- Claim Issue (Assessment) sub-aspect [Table 4 on item 5.3] consists of ‘Advance payment’, ‘Reimbursement denial’, ‘Concealment of medical records’, ‘Delayed service

response', 'Fax claim issues', 'Misapprehension' sub-aspect [Table 4 on item 5.3.1, 5.3.6] .

<u>Aspects identify on termination stage</u> are 'Impact level' (Dissatisfaction impact type) including sub-aspects that are identified under their aspects. [Table 3 on item 3.1 – 3.3, Table 4 on item 6]
Dissatisfaction impact type aspect has three sub-aspects. [Table 4 on item 6]

- High impact type sub-aspect [Table 4 on item 6.1] has three sub-aspects. [Table 4 on item 6.1.1-6.1.3]
- Medium impact type sub-aspect [Table 4 on item 6.2] has six sub-aspects. [Table 4 on item 6.2.1-6.2.6]
- Low impact type sub-aspect [Table 4 on item 6.3] has one sub-aspect. [Table 4 on item 6.3.1, 6.3.2]

<u>Aspects identify on all stage</u> are 'Social customer relationship management' [Table 4 on item 7' consists of 'Customer acquisition', 'Customer retention', 'Customer termination', 'Public relation' [Table 4 on item 7.1-7.4].

3. *Customer termination*

Customer termination includes cancellation, reject and termination action, occurs from some dissatisfaction took place or the period of contract reach to the maturity date. Exploring customers' dissatisfaction is one way to reduce a termination rate by find out misunderstanding issues from social media or make a questionnaire or implement a conservation system. [Table 3 on item 3.1 – 3.3]

Impact levels (Dissatisfaction types) that motivate customer termination are classified into three types.

3.1 High impact type such as customer make sue on the court, make a petition for OIC (Office of Insurance Commission is the organization which has a duty to control and promote Thai Insurance industry) or accuse agent. [Table 3 on item 3.1, Table 4 on item 6.1]

3.2 Medium impact type, such as make a policy cancellation or policy rejects. [Table 3 on item 3.2, Table 4 on item 6.2]

3.3 Low impact type includes giving negative feedback to the complaint system directly like a questionnaire or expressing a negative feeling of their experience on social media. [Table 3 on item 3.3, Table 4 on item 6.3]

<u>Aspects identify on termination stage</u> are 'Impact level' (Dissatisfaction impact type') including sub-aspects that are identified under their aspects. [Table 3 on item 3.1 – 3.3]

Aspects identify on termination stage are 'Impact level' including sub-aspects that are identified under their aspects. [Table 3 on item 3.1 – 3.3, Table 4 on item 6].

Dissatisfaction impact type aspect has three sub-aspects. [Table 4 on item 6]

- High impact type sub-aspect [Table 4 on item 6.1] has three sub-aspects. [Table 4 on item 6.1.1-6.1.3]
- Medium impact type sub-aspect [Table 4 on item 6.2] has six sub-aspects. [Table 4 on item 6.2.1-6.2.6]
- Low impact type sub-aspect [Table 4 on item 6.3] has one sub-aspect. [Table 4 on item 6.3.1]

1.2 Sentiment Extracted Tool - "Sansarn Tagging Tool"

This research, we used "***Sansarn Tagging tools***" is a tool for helping extract seed of words (keywords) which has designed and implemented for support this task using PHP programming. This tool provides the modules to pull out words that are under our designed aspects and sub-aspects [59]. We can extract three kinds of clue term (clue term) as our identified aspects/sub-aspects related to life insurance domain in Thai language.

The process starts from the crawling process to gather the user-generated contents from the criticized web-blogs named ***pantip.com***. This website is one of the famous critiqued websites in Thailand, which provides free blogs for users who participate in the Thai life insurance domain. They might be customers, insures, ex-customers, or prospects. People who have experience in life insurance express their sentiment via this website. This corpus from web-blogs consists of linguistics terms and non-linguistic terms such as Thai synonymous words, transliterated words, abbreviations, and slang. So the process of text preprocessing such sentence segmentation on corpus is required. The tagged corpus was prepared from random discussion topics and was annotated into each aspect and sub-aspects.

The main tagging modules are classified into five main modules which consist of '*Intention'*, *'Engagement Stage', 'Product (Topic)', 'Regulation (Topic)', Service Center (Topic)'*.

1. ***Intention*** means this tagging sentence shows the intention or purpose objective. It has five types of analysis that compose of
    a. Intention type means this tagging sentence shows requirement.
    b. Question type means this tagging sentence shows any suspicion by asking some question.
    c. Admiration type means this tagging sentence shows a positive feeling of the owner of the sentence.
    d. Complaint type means this tagging sentence shows a negative feeling of the owner of the sentence.
    e. Informative type means this tagging sentence gives some related information.
    f. Cause type this tagging sentence gives the reason for information.
    g. Other types.

Sentence (Chunk) *"รอเคลมนานมาก ประมาณ 4 ชั่วโมงได้ ได้ออกประมาณ 4 โมงเย็นแล้วต้องขนของหนีน้ำด้วย เลยจ่ายเองไปเลย / wait claim process too long time, wait around 4 hrs. go out at 16.00 pm and needs to take my belonging from flooding so we needs to paid by ourselves"*

- Clue word (Complaint-COM): *"รอเคลมนานมาก/wait claim process too long" and "เลยจ่ายเองไปเลย/so paid by ourselves"*
- Clue word (Informative-INF): *"ประมาณ 4 ชั่วโมงได้ ได้ออกประมาณ 4 โมงเย็น/wait around 4 hrs. go out at 16.00 pm.)"*

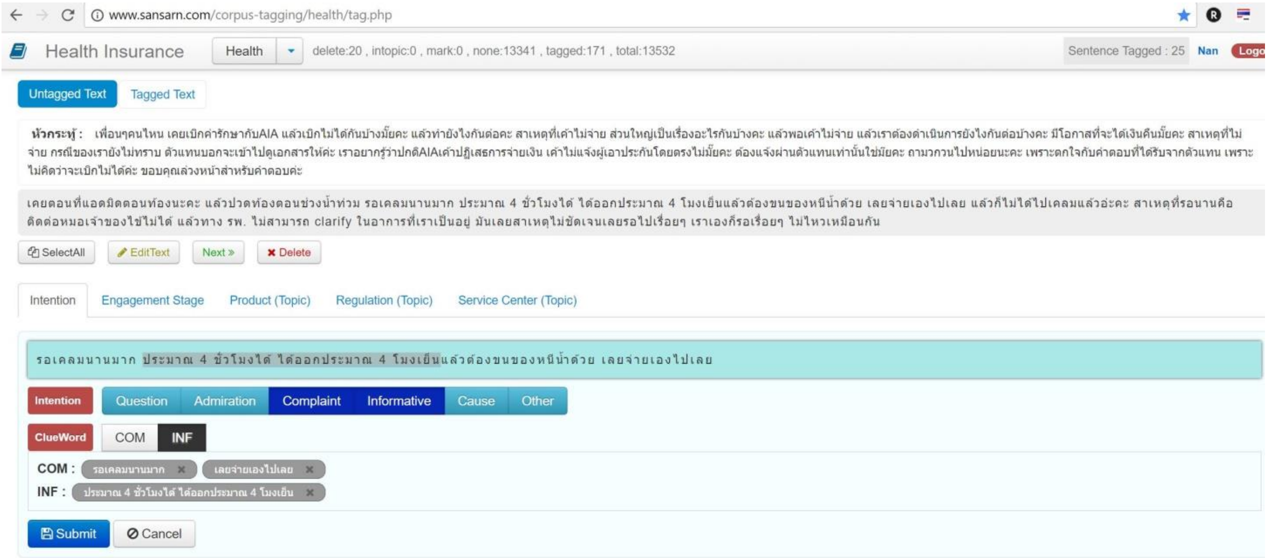


Figure 7: Sansarn tagging tools on ***Intention*** screen Example 1

Sentence (Chunk) “ที่บอกว่าประกันไม่ดี ประกันโกง นั่นนี่ มันมาจากตัวแทนที่เลวร้ายประพฤติตัวไร้จรรยาบรรณไม่กี่คน เท่าที่ผ่านมาก็เคลมได้ ไม่มีปัญหา/ that’s said, insurance is not good. Insurance cheat. It comes from a bad agent, behaved unethically, few people. In the past, claim process is no problem.”

- Clue word (Admiration-ADM): ‘เท่าที่ผ่านมาก็เคลมได้/ In the past, claim process is no problem’ and ‘ไม่มีปัญหา/no problem’
- Clue word (Complaint-COM): ‘ประกันไม่ดี/insurance is not good’, ‘ประกันโกง/ Insurance cheat’, ‘จากตัวแทนที่เลวร้าย/ a bad agent’ and ‘ประพฤติตัวไร้จรรยาบรรณ/ behaved unethically’

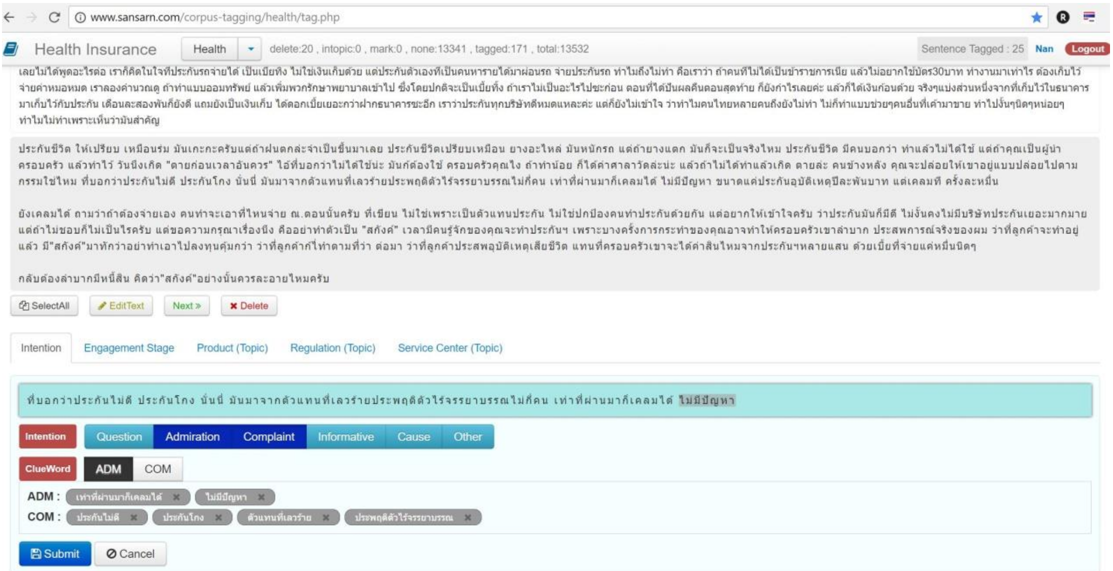


Figure 8: Sansarn tagging tools on ***Intention*** screen Example 2

2. ***Engagement Stage*** uses for specific the customer stage on customer life cycle such as prospect, customer or ex-customer, etc.
    a. Prospect stage
    b. Insured/Customer stage
    c. Ex-Insured/Ex-Customer stage
    d. Other

Sentence (Chunk) "ผมซื้ออยู่ ไม่ว่าจะ AIA เมืองไทย ไทยประกัน อยุธยาอลิอันซ์ ถ้าบริษัทไหนบอกว่าจ่ายเบี้ยเท่ากันไปตลอด บอกได้ 100% ว่าหลอกลวงครับ/ I bought, even if AIA, MuangThai, or Ayudhya Allianz the company says that pay the same amount to say 100% that is scam.."

- Clue word (Engagement Stage-INS): 'ผมซื้ออยู่/ I bought'

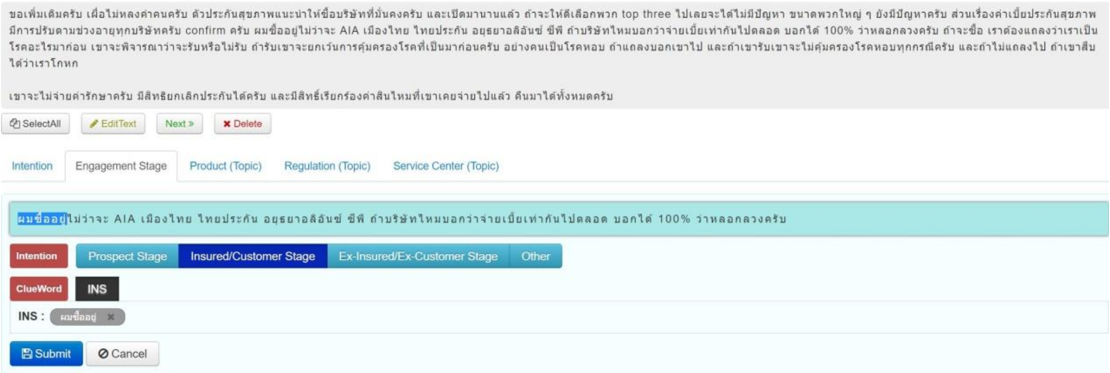


Figure 9: Sansarn tagging tools on ***Engagement*** screen

3. ***Product*** means this tagging sentence shows the concerned product. It has six groups of product that composes of
    a. Basic-Endowment
    b. Basic-Whole Life
    c. Basic-Term
    d. Basic-Universal Life
    e. Basic-Single License
    f. Rider
    g. Other

Sentence (Chunk) "ส่วนเรื่องค่าเบี้ยประกันสุขภาพ มีการปรับตามช่วงอายุทุกบริษัทครับ / Health insurance premium have to adjust to all ages range."

- Clue word (Engagement Stage-INS): 'ผมซื้ออยู่/ I bought'

4. ***Regulation*** means this tagging sentence shows the concerned operation and regulation. It has seven operation type ,which composes of
    a. Marketing Campaign
    b. Underwriting Rules
    c. Actuarial Rules
    d. Claim Process
    e. Policy Changing Rules
    f. Policy Operation
    g. Law
    h. Other

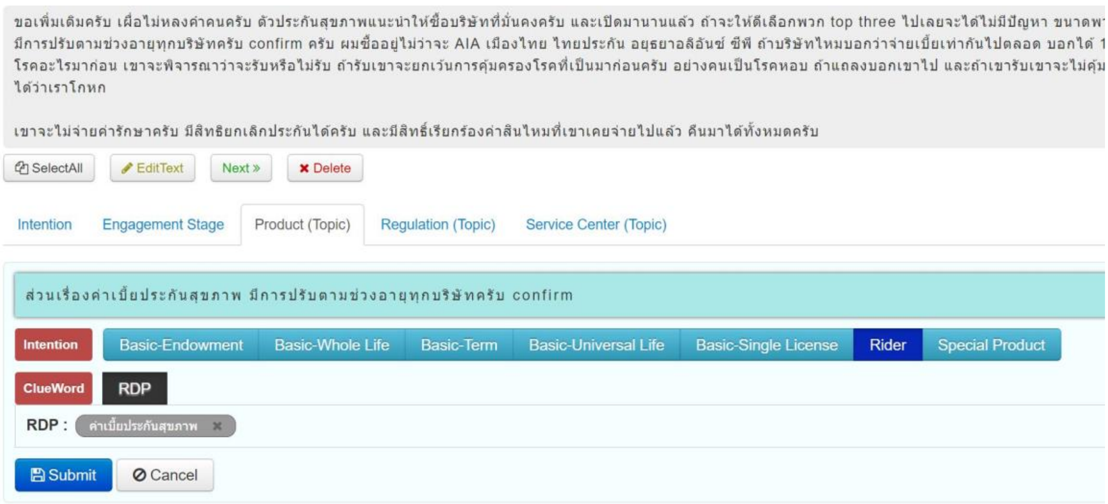


Figure 10: Sansarn tagging tools on ***Product*** screen

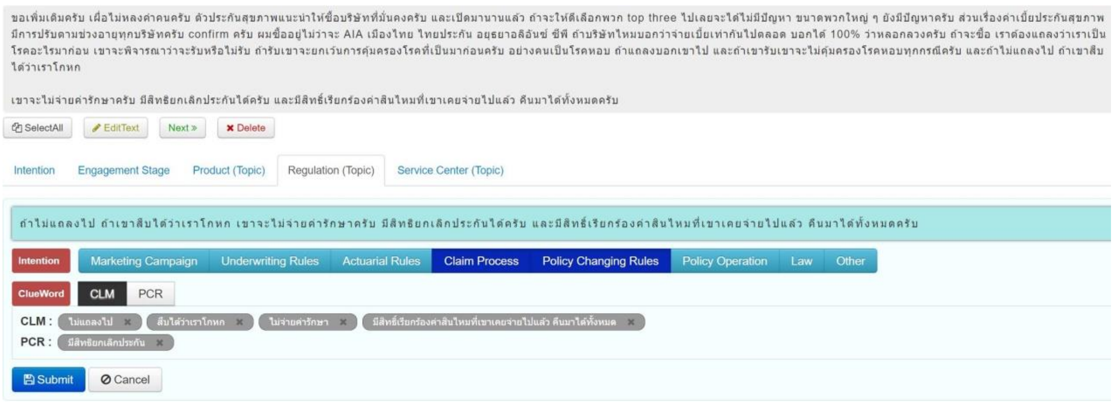


Figure 11: Sansarn tagging tools on ***Process*** screen 1

Sentence (Chunk) "ถ้าไม่แถลงไป ถ้าเขาสืบได้ว่าเราโกหก เขาจะไม่จ่ายค่ารักษาครับ มีสิทธิยกเลิกประกันได้ครับ และมีสิทธิเรียกร้องค่าสินไหมเขาเคยจ่ายไปแล้ว คืนมาได้ทั้งหมดครับ/ If concealing, if company know we are lying. The company can reject payment for treatment. Have the right to cancel the policy. And the company can request to return all of the expenses which the company paid for us. Return all."

- Clue word (Regulation (Topic)-CLM): 'ไม่แถลงไป / conceal,' 'สืบได้ว่าเราโกหก/ company know we are a lie,' 'ไม่จ่ายค่ารักษา/ reject payment for treatment,''มีสิทธิ์เรียกร้องค่าสินไหมเขาเคยจ่ายไปแล้ว คืนมาได้ทั้งหมด/ company can request to return all of the expense which company paid.'
- Clue word (Regulation (Topic)-PCR): 'มีสิทธิยกเลิกประกัน/ Have the right to cancel policy'

5. ***Service Center*** means this tagging sentence shows the concerned service touchpoint who handle this case. It has four service touchpoint types that composes of
    a. Agent and Agency
    b. Call Center

c. Customer Service Center
d. Training Center
e. Other

Sentence (Chunk) “ขึ้นกับตัวแทนจริงๆค่ะ เอ ไอ เอ ก็ดีนะคะ น้องสาวเราทำอยู่เคลมได้ดี และไม่มีเพิ่มเบี้ยค่ะ/, depend on agent, AIA is also good, my sister request claim assessment and process are good and no extra charge premium.”

- Clue word (Service Center (Topic)-AGT): ‘ตัวแทน/agent.’
- Clue word (Service Center (Topic)-SOT): ‘เอ ไอ เอ/AIA.’

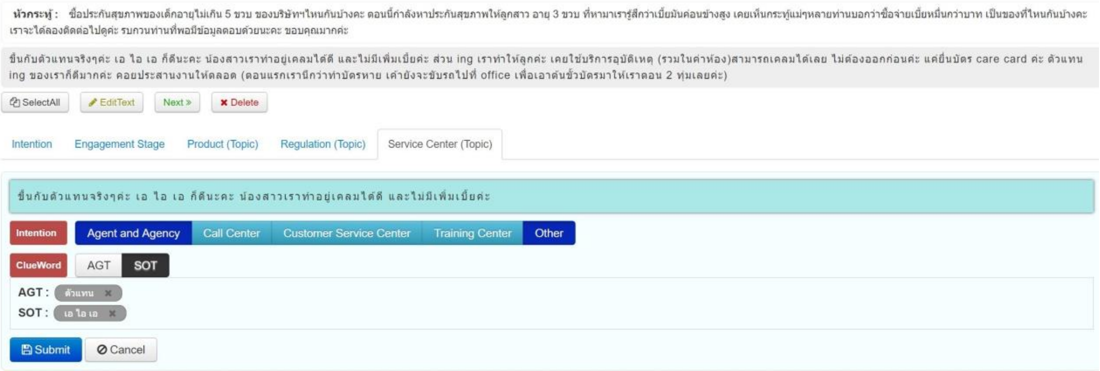


Figure 12: Sansarn tagging tools on ***Process*** screen 2

Developing programming and tools:

For data preprocessing, we use PHP (PHP Hypertext Preprocessor) and iMacro on Firefox to crawl data from websites and directly input it into Navicat for MySQL database version 8.0.8.

For text preprocessing, natural language processing use JAVA script and Python, and sentiment classification task use Javascript 1.7.
Tagging tools, SANSARN develop on PHP. For multidimensional sentiment cube generation, we develop on Microsoft SQL Server connect through Cognos Framework Manager 10.2, Cognos Transformer 10.2, Cognos Report builder 10.2.

Step 2. Extract and Validation

This is to define the knowledge structure, categorize the categories to search, restore, and use easily, improve, modify, or create some knowledge to be suitable for usage and analysis. This section presents our design of a multidimensional sentiment cube, and its dimension structure, which was created from our previous information extraction called sentiment extraction tool. Natural language processing, online analytical process (OLAP) and sentiment analysis concept in the aspect-based analysis was utilized in this task.

Design of Multidimensional Sentiment Cube

This section presents our design of a multidimensional sentiment cube and its dimension structure, followed by information extraction of sentiment keywords (Lexicon) and aspect-based sentiment analysis.

2.1. Design of Multidimensional Sentiment Cube and its Dimension Structure

A definition of a sentiment cube system as the previous section illustrates our proposed design of a multidimensional sentiment cube for the Thai life insurance business. It deployed the social CRM concept and its practice use with three main properties, i.e., multi-dimensions, hierarchical dimensions, and multi-measurements.

A sentiment cube system is proposed to store sentiment elements inside, together with their target aspects. Its data structure consists of a set of cells, organized in the form of dimension structure. The system keeps sentiment values associated with a set of entity features (aspects) and then populates a hierarchy of the cells in the sentiment cube with the sentiment values. An analysis is a process that aims to determine the attitude of a speaker concerning some topic. A dimensional scheme is a conceptual grouping of dimensions. It is a generalization of a constellation. Dimensions are grouped around nodes that model the dimensions that may be used together in the same analysis.

The automatic of sentiment analysis is the training process to identify sentiment within the contents. This multidimensional model is a conceptual model adapted to sentiment analysis from unstructured data. This model targets to provide the analyst with an adapted Online-Analytical Process (OLAP) conceptual view to manipulate the model's concepts using OLAP operations. The hierarchy of each dimension, including sentiment values associated with a set of entity features, and then populates a hierarchy of the cells in the sentiment cube with the sentiment values. Sentiment analysis modules affect a set of operations on the sentiment cube.

The cube generation needs to digest the contents from Thai life insurance critique blogs from narratives into the cube cell corresponding to each aspect (dimension – parent) and sub-aspect (sub-dimension – child) for analyzing the content across different cells, which related to the target dimension and measurement. This task gathers the data from critiqued posts, which extract to be various sub-aspects are, consists of tree-like hierarchically structured data. It will be clearly shown for easily navigation through data during analyses [84]. We derived the concept of the multidimensional cube, which is defined based on document base and increases a sentiment scheme for more potential explanations in a group feeling phenomenon on the definition as in the previous section [69].

In this research, we utilized the sentiment analysis concept to merge into a multidimensional cube to be a multidimensional sentiment cube.
We demonstrated in detail on our previous task on Thai Life Insurance business-related under social CRM concept.

Definition: **Multidimensional Sentiment Cube**

The formal description is as follows.

Let $SC = (D, H, S, F, M)$ be a multidimensional sentiment cube where $D = \{D_1, D_2, \ldots, D_n\}$ is a set of aspect and sentiment dimensions, $H = \{H_1, H_2, \ldots, H_n\}$ be the set of hierarchical structures. Each hierarchical structure H_i is a tree (a single-rooted directed acyclic graph), composed of a set of nodes (or vertices), and edges (or paths), $H_i = (N_i, E_i)$, where N_i is a set of nodes $\{n_{i0}, n_{i1}, n_{i2}, \ldots, n_{ipi}\}$ and E_i is a set of directed edges $\{e_{i0}, e_{i1}, e_{i2}, \ldots, e_{iqi}\}$, of the i-th dimension D_i.

Here, the apex (n_{i0}) of H_i represents the whole and n_{ij} N_i is an j-th instance node in the tree H_i. The sentiment set $S = \{+, -, 0\}$ expresses positive, negative and neutral sentiments. Each leaf node n_l is given a sentiment, denoted by $S(n_l)$. The footprint set $F = \{(h_{1k1}, h_{2k2}, \ldots, h_{nkn})\}$ is the set of all possible combinatorial footprints ($k_1, k_2, \ldots, k_n$) of H. The footprint-sentiment mapping M $F \times S \rightarrow$ R is the sentiment model, representing mapping of the footprints and sentiment types onto a real number (a membership value of the footprint in the specified sentiment type).

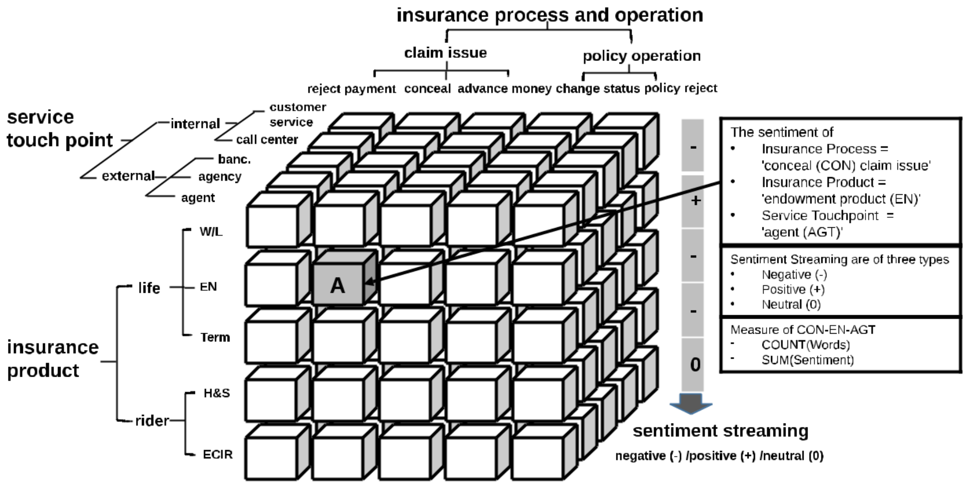


Figure 13: An example of a multidimensional sentiment cube for Thai life insurance business, where the x-axis expresses the insurance process and operation, the y-axis denotes the insurance product, and the z-axis represents the service touchpoint. When texts are streaming into the system, a sentiment (–, +, or 0) is extracted for each portion (sentence) in the texts, forming a stream of sentiment. The streaming sentiments are counted and kept in their corresponding cells.

Fig.13 shows an example of a sentiment cube, particularly designed for Thai life insurance business. The cube forms a multidimensional hierarchical platform wherein the cells are arranged in a structure of N-dimensional data, in this figure N = 3. Three dimensions define three entity aspects, i.e., 'insurance product (IP)', 'service touch-point (ST)', and 'insurance process and operation (IO)'. That is, D = {IP, ST, IO}. More concretely, the x-axis shows insurance process and operation, y-axis is an insurance product, and z-axis represents service touchpoint. Each cell stores a sentiment streaming value of negative, neutral, or positive (-, 0, or +).

Furthermore, an entity aspect can hold a set of sub-aspects (sub-dimensions), forming metadata that define dimensions and their associated sub-dimensions in a repeated manner. For example, the 'insurance product' has 'life' and 'rider' as its sub-aspect, and at the bottom layer, a sentiment value is stored. In the figure, the cell labeled with A is the sentiment of 'conceal (CON)' under claim issue of 'endowment product (EN)' when the service touchpoint is 'agent (AGT).' A sentiment storage module manages sentiment values associated with entity aspects (or features). Towards the construction of a sentiment cube, one needs to digest and summarize the contents in Thai life insurance critique blogs, often narratives, into the structure of a data cube. This sentiment value is derived from the summation of sentiment scores obtained from all textual fragments that relate to the focused topic, in this case, CON-EN-AGT. After the cube construction, we can use a well-defined set of online analytical processing (OLAP) operations to explore the complicated facts in the multidimensional and the hierarchical point of view.

With this OLAP concept, one can get an insight via navigation through data with exploratory analysis. A cell in a data cube typically corresponds to an aspect (dimension or parent) or a drill-down sub-aspect (sub-dimension or child) for analyzing the content across different cells related to the target dimension and measurement.

An example of insurance-related sentiment cubes

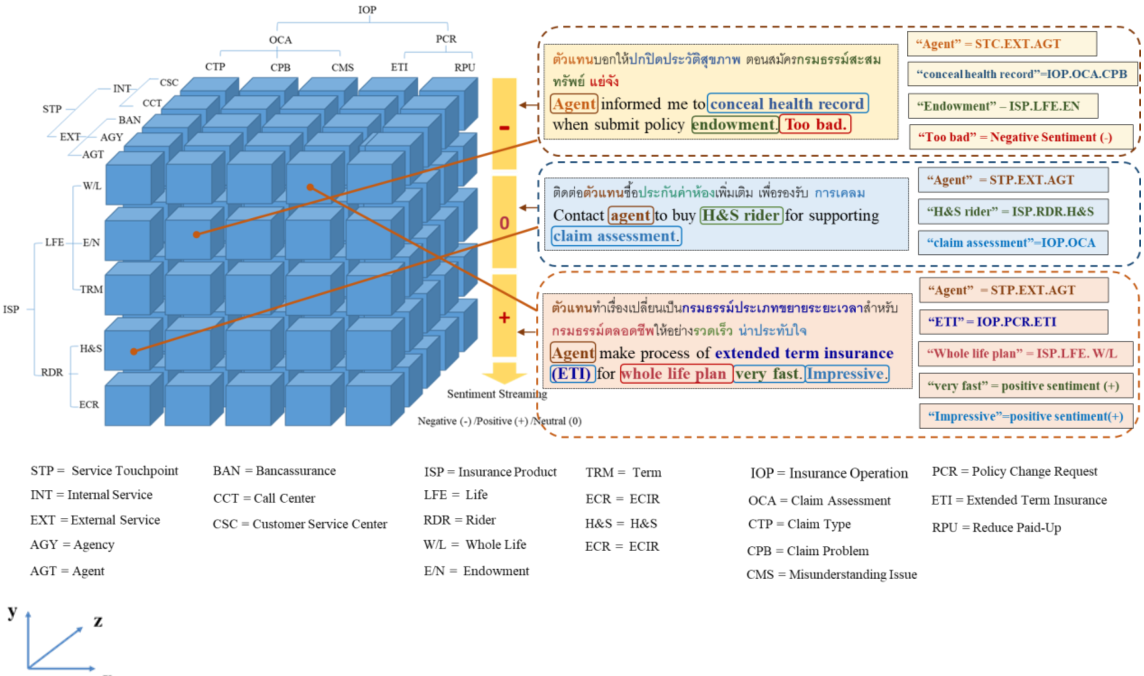


Figure 14: An example of a multidimensional sentiment cube for Thai life insurance business, where x-axis is issue characteristics, y-axis is product/service type, and z-axis is service touchpoint. Each cell stores a sentiment streaming value of -, 0, or +

Refer to Figure 14, assume that an insurance company designs a sentiment cube SC comprised of three dimensions. In this sentimental cube, three dimensions are expressed under three sentimental polarities; positive (+), negative (–), and neutral (0). In the figure, the x axis expresses issue type, the y axis denotes product/service type, and the z axis represents service touchpoint type.

$D = \{IOP, STP, ISP\}$ *(IOP = insurance operation, STP = service touchpoint, ISP = insurance product).*

$H_1 = \{IOP, \{OCA, PCR\}, \{(CTP, CPB, CMS), (ETI, RPU)\}$,

$H_2 = \{STP, \{INT, EXT\}, \{(AGY, AGP), (BAN, CCT, CSC)\}\}$, *and*

$H_3 = \{ISP, \{LFE, RDR\}, \{(W/L, E/N, TRM), (H\&S, ECR)\}\}$,

F = {(OCA, INT, AGT), (OCA, EXT, AGY), (OCA, BAN, CCT), (CTP, AGP, AGT), …} and M = {(OCA, INT, AGT, +):2.5, (OCA, EXT, AGY, –):4.5 , (OCA, BAN, CCT, 0):6, (CTP, AGP, AGT, +):0.2, …}.

Our task, generating the cube, needs to digest the contents from Thai life insurance critique blogs from narratives into the cube cell corresponding to each aspect (dimension – parent) and sub-aspect (sub-dimension – child) for analyzing the contents across different cells related to the target dimension and measurement. The example illustrations have to provide the following functions in four steps.

First, to identify the aspect (dimension) and sub-aspect (sub-dimension) in a hierarchical structure. The second is to make a process of data preparation, text mining process with lexicon preparation, and mapping unstructured data corresponding with the designate as aspect (dimension) and sub-aspect (sub-dimension) related to measurements determination. The third is to design three cube model generations, and the last is to make an efficient materialization.

However, our research utilizes a multidimensional cube with sentiment analysis called 'multidimensional sentiment cube.' The measurements consist of the number of words or phrases matching with the lexicon and their sentiment score in each dimension.

2.2 Dimension structure design on aspect identification from business understanding

The design of a multidimensional sentiment cube for the Thai life insurance business using social CRM and its practice starts from declared three main properties (multi-dimensions, hierarchical dimensions, and multi-measurements).
Multi-dimensions are gathered from the individual dimension, assigned from each aspect under covered by analysis criteria. To identify aspects and sub-aspect under the form of hierarchical structure, which are the main elements in cube generation, requires a business understanding of its principle. Hierarchical is a many-to-one relationship on specific levels, representing a relationship among different sub-aspects under the same aspects within a hierarchy [35]. Therefore hierarchical dimensions are represented by all of the sub-aspects in different layers under the same aspect relationship. When designing the group of aspects to be multi-dimensions, it needs to concern with hierarchical dimensions at the same time.

For easy understanding, we draw a process flow on **business process mining notation (BPMN)** for process reference as below.

A line area shows the process of approaching new customers and submits new policy contract. It indicates that the processes which effect on acquisition stage are: the process of product (plan) offering to the customer, the process of policy approval, and issue of a new policy.

B line area shows the process to announce a result of policy approval. The process of policy approval so this process checking step is checking (prepare new policy contract / inform result accept/reject to a customer).

C line area shows the process to make a request on the claim assessment process. Customers require after-sales service.

D line area shows the claim verification process. In general, insured expect to get compensations cover all of their payment.

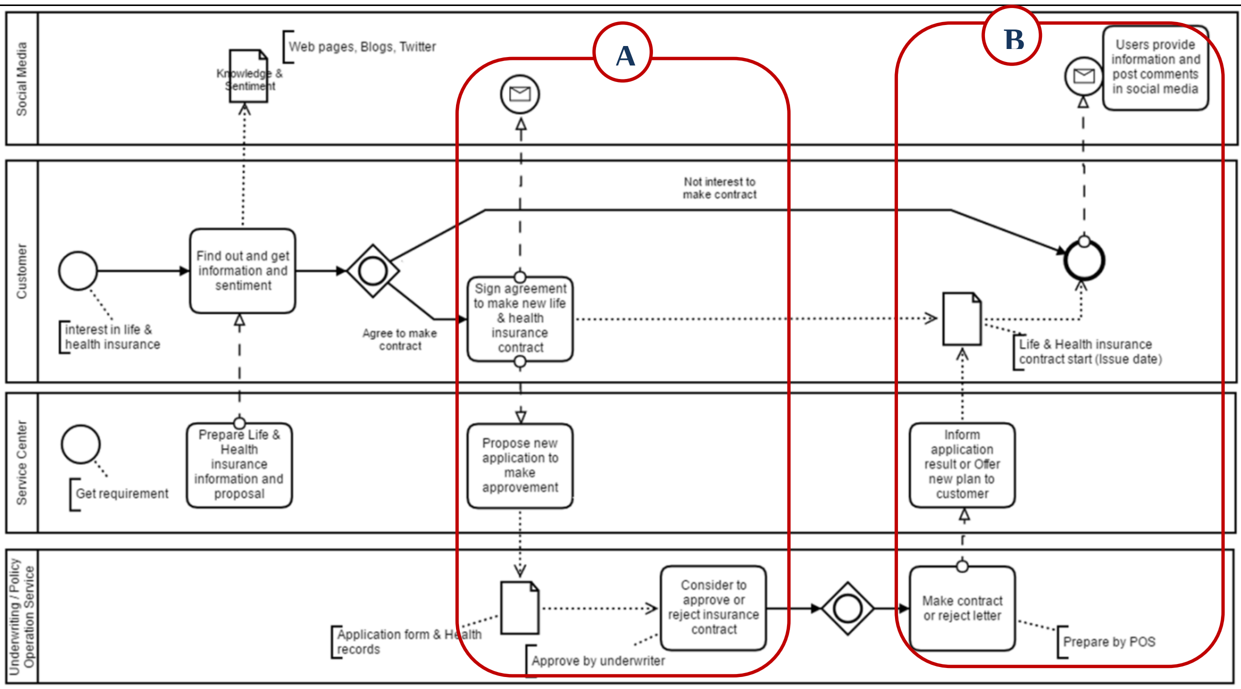


Figure 15: Customer Acquisition Stage in Life Insurance with line A and B

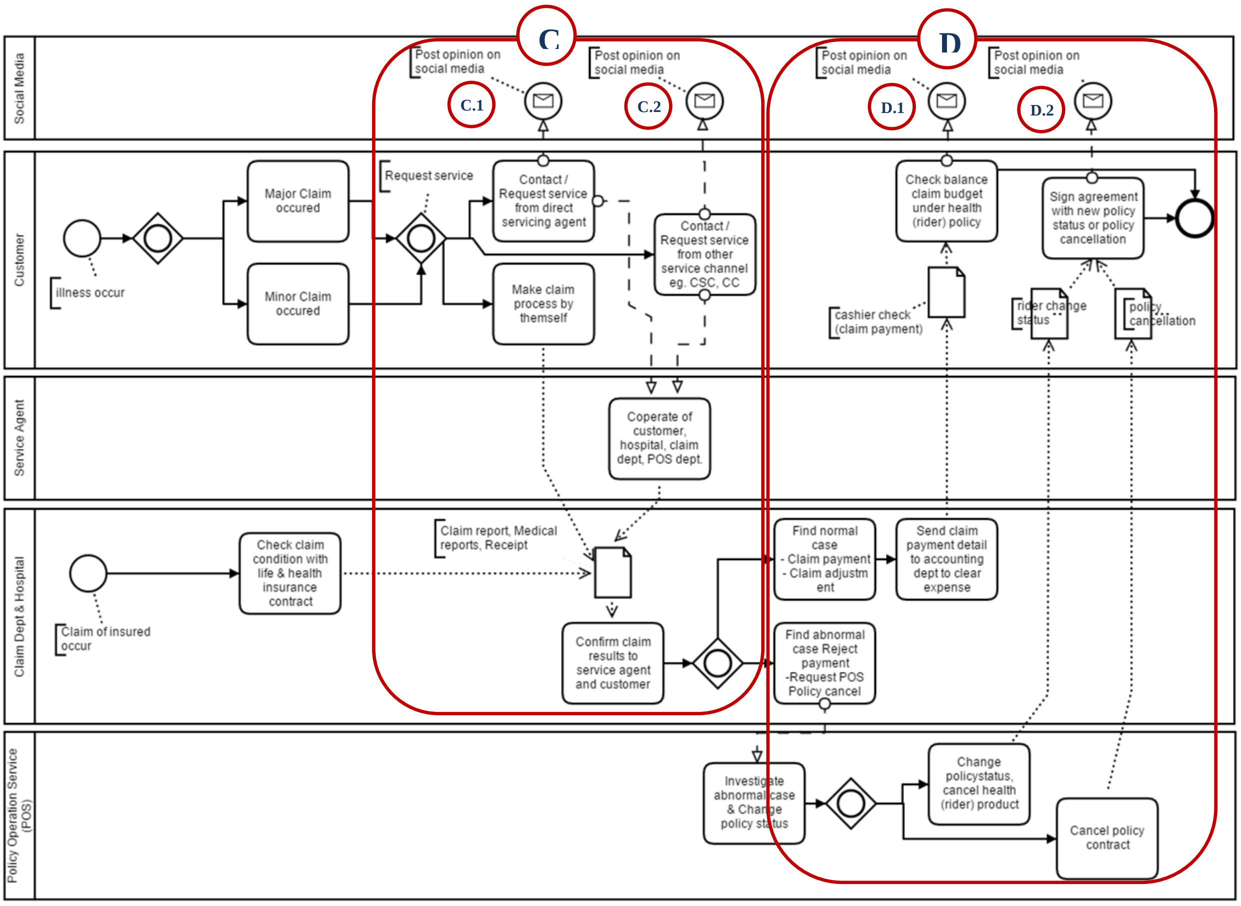


Figure 16: Customer Engagement & Retention Stage in Life Insurance

Aspects and Sub-aspects Identification

Our task introduces two kinds of aspect which have to concern dimension structure design.

a. Ownership representative aspect is the owner of subject or object which was expressed sentiment or mentioned by opinion reviewer under analysis purpose of multidimensional sentiment cube.

b. Analysis viewpoint aspect is subject or objects which were expressed sentiment or mentioned by opinion reviewer under analysis purpose of multidimensional sentiment cube.

Ownership representative aspect:

The purpose of our multidimensional sentiment cube is under a business viewpoint, so it is indispensable to state about company aspect and product aspect (plan in life insurance).

A 'Company' aspect of identification on representative ownership aspect is the Thai life insurance company name. We gathered user-generated contents from various companies such as American International Assurance (AIA), Thai Life Insurance, Muang Thai Life Assurance, Allianz Ayudhya, Bangkok Life Assurance, Thai Samsung Life Insurance, SCBLIFE Assurance, including with some health insurance such as BUPA health insurance, SIGNA health insurance, etc. [Table 4 on item 1]

'Product aspect identification on representative ownership aspect is 'Life' (Basic plan) and 'Rider' (Supplemental plan). Life is the main product consists of Endowment, Whole Life, Investment, Retirement and Term [Table 4 on item 2.1.1-2.1.5] and Rider is supplement product composes of HS, HB, AI, AI/RCC, ECIR, CR, disability and WP. [Table 4 on item 2.2.1-2.2.8]

Analysis viewpoint aspect:

Analysis viewpoint aspect identification starts from analyzing the relationship among business functions which impact on customer sentiment. This step explains the relation of life insurance functions to the CRM concept of service process that their negative or positive sentiment can use to improve customer satisfaction. After finding out their relationship among main aspects, the system eventually presents the illustration of main properties; multi-dimensions, hierarchical dimensions, and multi-measurements.

Aspects identify on acquisition stage are 'Service touchpoint' and 'Service evaluation factor' including sub-aspects that are identified under their aspects. [Table 3 on item 1.1 – 1.5, Table 4 on item 3&4]

Service touchpoint aspect has two sub-aspects. [Table 4 on item 3]

- Internal service sub-aspect [Table 4 on item 3.1] has three sub-aspects. [Table 4 on item 3.1.1-3.1.3]
- External service sub-aspect [Table 4 on item 3.2] has five sub-aspects. [Table 4 on item 4, 4.1-4.5]

Service evaluation factor aspect has five sub-aspects. [Table 4 on item 4, 4.1-4.5]

Aspects identify on intention stage is 'Process and Operation' including sub-aspects also identify under their aspects. [Table 3 on item 2.1 – 2.3, Table 4 on item 5]

Process and Operation aspect has three sub-aspects. [Table 4 on item 5]

- Policy Operation sub-aspect [Table 4 on item 5.1]
- Claim Operation (type) sub-aspect [Table 4 on item 5.2] has two sub-aspects. [Table 4 on item 5.2.1-5.2.2]
- Claim Issue (Assessment) sub-aspect [Table 4 on item 5.3] consists of 'Advance payment', 'Reimbursement denial', 'Concealment of medical records', 'Delayed service response', 'Fax claim issues', 'Misapprehension' sub-aspect [Table 4 on item 5.3.1, 5.3.6] .

<u>Aspects identify on termination stage</u> are 'Impact level' (Dissatisfaction impact type) including sub-aspects that are identified under their aspects. [Table 3 on item 3.1 – 3.3, Table 4 on item 6]

Dissatisfaction impact type aspect has three sub-aspects. [Table 4 on item 6]

- High impact type sub-aspect [Table 4 on item 6.1] has three sub-aspects. [Table 4 on item 6.1.1-6.1.3]
- Medium impact type sub-aspect [Table 4 on item 6.2] has six sub-aspects. [Table 4 on item 6.2.1-6.2.6]
- Low impact type sub-aspect [Table 4 on item 6.3] has one sub-aspect. [Table 4 on item 6.3.1, 6.3.2]

<u>Aspects identify on all stage</u> are 'Social customer relationship management' [Table 4 on item 7' consists of 'Customer acquisition', 'Customer retention', 'Customer termination', 'Public relation' [Table 4 on item 7.1-7.4].

The identification of Aspect & Sub-aspect summarization: (Multi-Dimensions Determining)

No.	(Dimension/Sub-dimension)	No.	(Dimension/Sub-dimension)
0.	**Transaction ID**	**5.**	**Process and operation**
1.	**Company X**		5.1 Policy operation
	1.1 Company X		5.2 Claim operation (Claim type)
	1.2 Company Y		5.2.1 Minor claim
	1.3 Others (No information)		5.2.2 Major claim
2.	**Product**		5.3 Claim Issue (Claim assessment)
	2.1 Life (Basic)		5.3.1 Advance payment
	2.1.1 Endowment		5.3.2 Reimbursement denial
	2.1.2 Whole life		5.3.3 Concealment of medical records
	2.1.3 Investment		5.3.4 Delayed service response
	2.1.4 Retirement		5.3.5 Fax claim issues
	2.1.5 Term		5.3.6 Misapprehension
	2.2 Rider (Supplement)		5.3.6.1 Misapprehension in coverage
	2.2.1 Hospital & Surgical expenses (H&S)		5.3.6.2 Misapprehension in protection
	2.2.2 Hospital benefit (HB)		5.3.6.3 Misapprehension in period
	2.2.3 Accident indemnity rider (AI)		5.3.6.4 Misapprehension in exception
	2.2.4 AI/riot&civil commotion (AI/RCC)	**6.**	**Impact level**
	2.2.5 Enhanced critical illness rider (ECIR)		6.1 High impact
	2.2.6 Cancer death & income benefit (CR)		6.1.1 Lawsuit
	2.2.7 Disability		6.1.2 Petition to off. of insurance com.(OIC)
	2.2.8 Waive premium (WP)		6.1.3 Make an accusation to agent
	2.3 Affinity marketing		6.2 Medium impact
	2.4 Special product		6.2.1 Policy cancellation
	2.4.1 Social insurance		6.2.2 Policy rejection

	2.4.2 Thai 30-BHT health insurance		6.2.3 Policy invalid
	2.4 Others (No information)		6.2.4 Policy void
3.	**Service touchpoint**		6.2.5 Policy surrender
	3.1 Internal		6.2.6 Loan
	3.1.1 Call center & customer service center		6.3 Low impact
	3.1.2 Operation dept.: claim, underwriting		6.3.1 Negative impression
	3.1.3 Telemarketing		6.3.2 Positive impression
	3.2 External	**7.**	**Social Customer relation management**
	3.2.1 Agent or agency		7.1 Customer acquisition
	3.2.2 Bancassurance		7.2 Customer retention
	3.2.3 Credit card		7.3 Customer termination
	3.3 Others		7.4 Public relations
4.	**Service Characteristics (Service evaluation)**		
	4.1 Reliability		
	4.2 Responsiveness		
	4.3 Characteristics		
	4.4 Consistency		
	4.5 Knowledgeability		
	4.6 Others		

8.	**Measurement**	
	8.1 No. of sentences	8.10 No. of words on process & operation
	8.2 No. of words on company	8.11 Total scores on process & operation
	8.3 Total scores on company	8.12 No. of words on impact level
	8.4 No. of words on product	8.13 Total scores on impact level
	8.5 Total scores on product	8.14 No. of words on negative words
	8.6 No. of Words on service touchpoint	8.15 Total scores on negative words
	8.7 Total Scores on service touchpoint	8.16 No. of words on positive words
	8.8 No. of words on service characteristic	8.17 Total scores on positive words
	8.9 Total scores on service characteristic	8.18 Total scores on service characteristic

Table 4: List of dimensions, sub-dimensions and measurements of multidimensional sentiment cube

From business knowledge understanding as above, we analyzed the relationship among business functions which impact on customer sentiment. Then we summarize the multi-aspects which can refer to be dimension of multidimensional sentiment cube as table 4. Multi-dimensions and hierarchy dimensions design depend on dimensions and sub-dimensions design. Dimension is aspect and sub-dimension is sub-aspect. Table 4 shows main item is aspect (dimension) such as line no. 1, 2, 3, 4, 5, 6. Sub-aspect (sub-dimension) is the next column layer such as 1.1, 1.2, 2.1, 2.2, 3.1, 3.2, etc.

However, sub-aspect is possible occur the multi-layers. It depends on the analysis requirement so next of next layer of table 4 such as 2.1.1, 2.1.2, 2.1.1 ... will refer to sub-aspect as well. This characteristic is similar to hierarchical structure. Therefore aspect (dimension) is parent level and sub-aspect is child level in the hierarchical structure of multidimensional sentiment cube likewise.

2.3 Methodology of Multi-dimensional sentiment cube

A recommendation method using multidimensional sentiment mining by taking benefit of multidimensional sentiment cube and data mining method as following: (Figure 17)

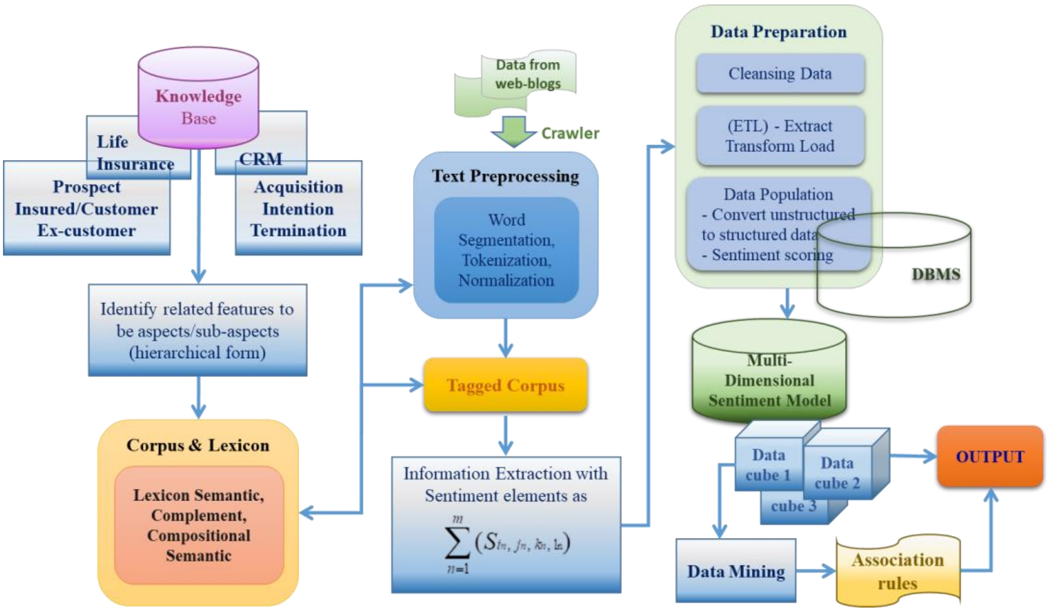


Figure 17: Multidimensional sentiment cube mining methodology

1. *Knowledge Base preparation:* Study business knowledge and preparation the relationship of features in our case study. This process is possible to make participation with domain expert. (Refer to Item 3.1) The hierarchical structures based on business knowledge are as flow below.

2. *Identify related features in hierarchical form:*

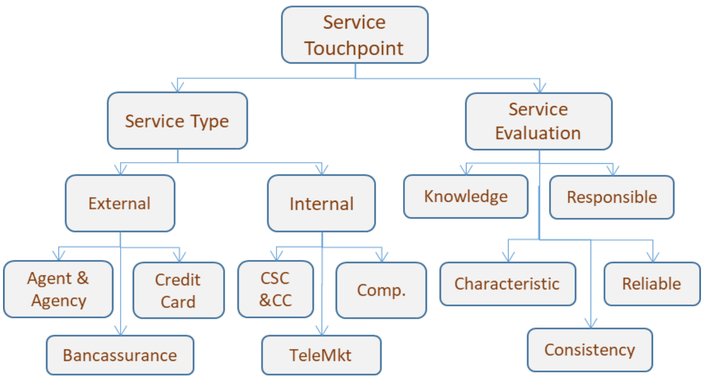


Figure 18. Service Touchpoint Hierarchy

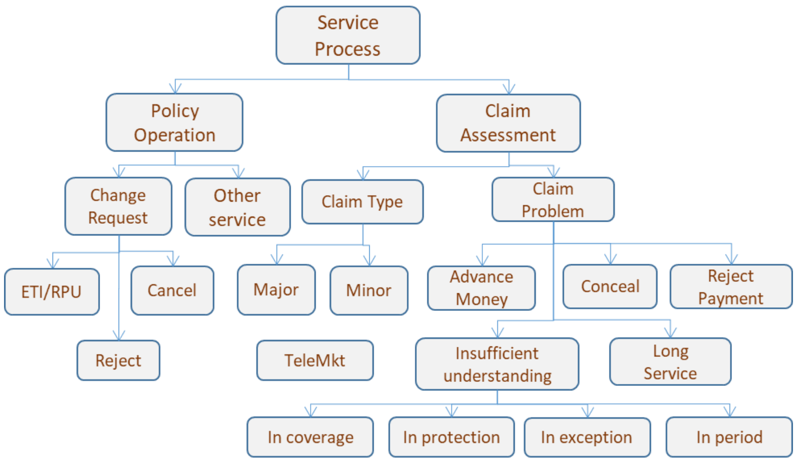


Figure 19. Service Process: Policy operation & Claim assessment Hierarchy

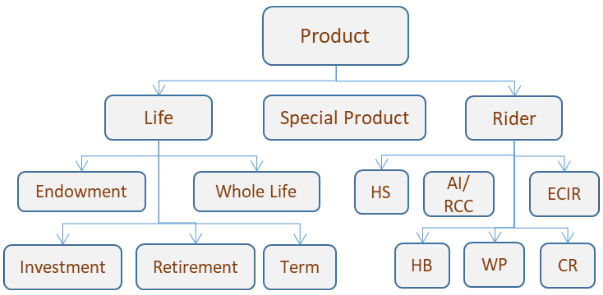


Figure 20. Product (Plan) Hierarchy

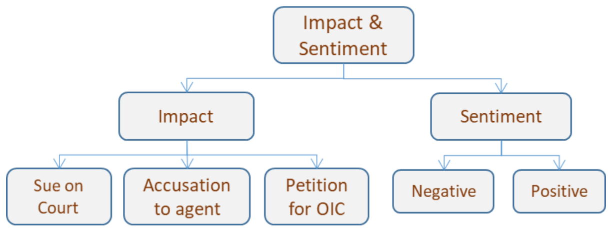


Figure 21. Impact Type Hierarchy

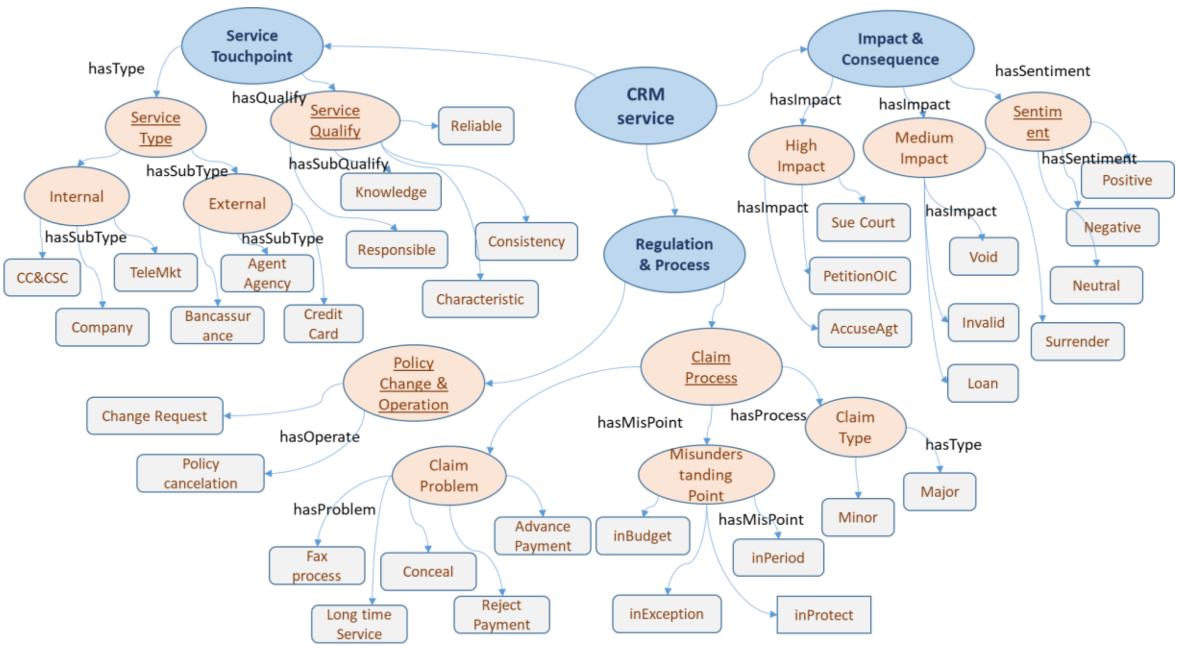


Figure 22. Related Aspects Hierarchy

3. Corpus and Lexicon construction based on aspect-based sentiment analysis:

Aspect-Based sentiment analysis concept:

We deploy the concept of sentiment analysis for extracting sentiment and aspects elements from unstructured data by using standard opinion are a quintuple; [24, 28].

$$(a_i, sa_{ij}, s_{ijkl}, r_k, p_l)$$

Where a_i is an aspect name, sa_{ij} is a sub-aspect of a_i, s_{ijkl} is the sentiment on sub-aspect sa_{ij} of aspect a_i, r_k is the opinion reviewer, and p_l is the period when the opinion is expressed by r_k. The sentiment s_{ijkl} is positive, negative, or neutral

This research require total sentiment of customer satisfaction, so

Total score of customers' sentiment = $\sum_{n=1}^{m} (S_{i_n, j_n, k_n, l_n})$

Whereas $S_{in, jn, kn, \ln} \in (a_{i_n}, sa_{i_n, j_n}, S_{i_n, j_n, k_n, l_n}, r_{k_n})$

Where a_i is an aspect name, sa_{ij} is a sub-aspect of a_i, s_{ijkl} is the sentiment on sub-aspect sa_{ij} of aspect a_i, r_k is the opinion reviewer.
The sentiment s_{ijkl} is positive, negative and n is number of critiqued word chunk.

When a_i = {prd, sp, is} is a set of aspects (parent dimension of a cube)
r_k = {st, cm} is a set of opinion reviewer (parent dimension of a cube)
$sa_{in} \in (sa11_{ij}, sal2_{ij})$

$sal1_{ij} = \{lf_{prd}, rd_{prd}, am_{st}, sp_{st}, po_{sp}, stt_{st}, sv_{st}, po_{sp}, ca_{sp}, hi_{is}, mi_{is}, si_{is}\}$ is a set of sub-aspects (child dimension of a cube level1)

$sal2ij = \{ct_{ca}, cp_{ca}\}$ is a set of sub-aspects (child dimension of a cube level2)
s_{ijkl} = {negative, positive}

Aspect in parent level: st = service touchpoint, *cm* = company, *prd* = product(plan), *sp* = service process, *is* = impact & sentiment

Sub-aspect in child level 1:
lf_{prd} = life (Basic) of *prd*, rd_{prd} = rider of *prd*,
am_{prd} = affinity marketing of *prd*, sp_{prd} = special product of *prd*,
stt_{st} = service type of *st*, sv_{st} = service evaluation of *st*,
po_{sp} = service operation of *sp*, ca_{sp} = claim assessment of *sp*
hi_{is} = high impact type of is, mi_{is} = medium impact type of *is*,
si_{is} = low impact type of *is*

Sub-aspect in child level 2:

ct_{ca} = claim type of *sp*, cp_{ca} = claim problem of *sp*

4. *Text preprocessing and Information extraction with Text mining and Natural Language Processing (NLP) in Thai language.*

This section has many tasks which need to process respectively below:

- Create lexicon construction using aspect-based sentiment analysis concept
    - o Refer to Section 3.2, Aspect-based sentiment analysis concept with formula ➔ Define : an aspect name as a_i, a sub-aspect of a_i as sa_{ij}, sentiment on sub-aspect sa_{ij} of aspect a_i as s_{ijkl}, r_k is the opinion reviewer
    - o Text preprocessing in Thai language with text mining and NLP
        - Corpus preparation – www.pantip.com
        - Lexicon construction as specific aspect/sub-aspects with "SANSARN tagging tool"
        - Text preprocessing process
    - o Define sentiment word scoring by SO-CAL concept
- Exploiting linguistic knowledge for sentiment classification
- Information extraction process from corpus and lexicon => convert unstructured data to structure data
- Database management (ETL,…cleansing)
- Data population in database – Fact table

The significance of lexicon construction for multidimensional sentiment cube on unstructured data with text mining:

Text mining is the discovering process of new knowledge through the analysis of the text. Information extraction task is one of the main tasks of the text mining process to dig out the words or phrases that match the related lexicon from the main corpus. Information extraction also analyzes unstructured text and identifies key phrases and relationships within the text [18]. To analyze unstructured data to get hidden knowledge from all data bring more advantages to a business. However, traditional cube dealing well with structured data, so there are challenges for analyzing unstructured data into cube merge with sentiment analysis. Unstructured data such as narrative text fields bury causal factors which do not explicitly exist. It is important to flexibly support analysts in that data mining with OLAP and text content analysis (Unstructured data) in an integrative manner [99]. However, OLAP can only support such integrative analysis. OLAP can support drill-down and roll-up feature on structured attributed dimensions. However, they cannot analyze support drill-down or roll-up on unstructured data such as critiques sentences from people.

For this reason, understanding the analysis viewpoint requires knowledge to understand the cube relate to the contents of all the narratives corresponding with their measurements. The text mining in this task has the main purpose of building a particular lexicon that supports CRM and life insurance business knowledge.

The significant of NLP on text mining and aspect-based sentiment analysis:

In general, non-standard language, such as the Asian language, has unique characteristics, especially, clarify the end of each sentence is difficult. The Thai language also has a lot of ambiguous meaning and conversational language in sentences. The Thai language's challenge problems are how to manage the long sentence and how to separate wording with correct meaning [16, 61, 63, 77]. To extract sentiment from the Thai language with a standard dictionary or shared corpus, the same as the English language, is not easy. Especially, word of the mount from social media is a free-form format that consists of chat language, slang, or transliterated words, etc. It requires a text preprocessing process with NLP to handle it.

To concern about cube usage, a hierarchical platform need to be recognized. Lexicon characteristics should be declared words, phrases, or sentiment words for extracting at the lowest level of a cell of sub-aspects, which is able to support drill-down and roll-up of cube functions. Well-identified sentiment words of the lexicon contribute to the accuracy of sentiment analysis. To support the drill-down or roll-up of a cube, the aspect, and sub-aspect hierarchical structure is needed to define such a tree diagram. In real-world applications, the polarities of many words depend on the aspects [45], and the differentiation of prior and contextual polarity is crucial [83]. For example, "มีเซอร์วิสมายด์" under service means "(someone) has a service mind." This word does typically not have any characters to show positive meaning. However, for the service domain, this word has a positive meaning. So text preprocessing process for a special lexicon is required. Data preprocessing task is important to examine the effectiveness and efficiency of customer analytics in order to make the right decisions in producing new products or services based on customer satisfaction.

Corpus preparation and lexicon preparation is the main task for extracting the elements of aspects, sub-aspects, and sentiment from unstructured data and then rearranging them into a structured platform, which will be treated as an attribute dimension of the cube. To transform unstructured data to be structured, data require some principle for defining a definition of targets. Our task will extract aspects, sub-aspects, and sentiment, using the total score of customers 'sentiment definition as above.

4.1. Corpus preparation

The text preprocessing process starts from the crawling procedure to gather user-generated content from the criticized web-blogs named ***pantip.com,*** one of the famous critiqued websites in Thailand providing free blogs for users who participate in Thai life insurance. This corpus from web-blogs consists of linguistics terms and non-linguistic terms such as Thai synonymous words, transliterated words, abbreviations, and slang. So the process of text preprocessing such sentence segmentation on corpus is required.
Collections of the corpus have different volume data. We made the equal volume of data by preparing a unit group of words on sentence segmentation under conditions. One divided sentence contains at least one subject and one verb. Divided sentences are separated from each users' message in each topic using space and determined length of Thai characters per line around 200-300 characters. The tagged corpus was prepared by sentiment extracted tool, 'Sansarn tagging tool' from the previous section. Next process, we did word segmentation, tokenization, and normalization process and then extract key feature terms of phrases from a given text as Figure 14.

4.2. Lexicon Construction

The design and completeness of the lexicon base affect the performance of the sentiment analysis. Lexicon is the terms of a word which found the words matched in ***LEXiTRON,*** Thai-English electronic dictionary provided by NECTEC [56]. The first kind of lexicon we use to track matching words is similar to a standard dictionary in general meaning. To support a special word's meaning-extraction, it requires preparing another kind of lexicon called clue term. They were prepared for special words, phrases, or groups of words that cannot match precisely in *LEXiTRON* for identifying the group of aspects/sub-aspects, including sentiment words. For example, the "*service mind*" word in the service domain shows *positive* sentiment on the service touchpoint aspect dimension, "*fax claim*" shows the claim aspect dimension about fax document but does not have any senses.

Clue term lexicon preparation in this task divided the lexicons into three types:

a.) Lexicon semantics clue Terms:

We arranged this type of clue term by the semantic meaning of ***linguistic rules*** in Thai syntax using a noun, verb, and particular term patterns. Besides, we added ***non-linguistic words*** such as slang words, ambiguous words, and transliterated words.

b.) Complement clue Terms:

For improving the sentences' semantic, we classified the standard term in many types of the word, including linguistic and non-linguistic platforms. Linguistic terms contain abstract noun (*การ, ความ/~ing*), preposition (*ในช่วง เวลา/during*, *เกี่ยวกับ/about*, *ในระหว่างนั้น/between*), auxiliary verb (*ควรจะ/should, ต้อง/must*), general adverb (*วิเศษณ์, สุดยอด/excellently, แย่/terrible*) conjunction (*ดังนั้น/therefore, อย่างไรก็ตาม/however, มิฉะนั้น/otherwise*) and the other kind of noun. All of them are terms that do not have any meaning relate to each analysis group.

c.) Compositional semantics clue terms

These clue terms have a meaning related to pragmatics in words and phrase level. They consist of technical vocabulary related to life insurance such as *การสิ้นสุดประกันภัย/Lapse, การประกันภัยขยายเวลา/Extended Terms Insurance, สถานะมีผลบังคับของกรมธรรม์/Inforce, การเรียกร้องให้จ่ายเงินตามสิทธิ์ในกรมธรรม์ -เคลม/Claim assesment, กรมธรรม์ใช้เงินสำเร็จ/Paid-up policy*, etc.

Besides, we gathered phrases that were negatively or positively felt when expressed in the insurance or service domain. This word's sense depends on the story, and sometimes they do not have any feelings when they are in a single word. For example, *ไม่มีจรรยาบรรณ/No ethic*, *ไม่ได้แถลงข้อมูลอันเป็นข้อเท็จจริง/Did not declare the factual information, ยกเลิกกรมธรรม์/Reject policy, ปฏิเสธการจ่ายสินไหม/Reject payment*, etc.

Clue word							
Lexicon :	Word Type	Intention			Topic		
		Question	Sentiment		Product	Regulation	Service Center
			Positive	Negative			
Linguistic word	common noun	คำถาม/questions,ข้อสงสัย/ doubt, วิธีคำนวณเบี้ยประกัน insurance premium calculation	ความดีใจ/gladness, ความยินดี, ความเบิกบาน/ cheerfulness	ความเสียใจ/ disappointment, ความเดือดร้อน/ affliction	ประกันแบบสะสมทรัพย์/endowment, ประกันแบบชั่วคราว/ term, ประกันแบบตลอดชีพ/whole life	เรื่องร้องเรียน/ complaint,ค่ารักษาพยาบาล, medical expense,กรมการประกันภัย/department of insurance	ตัวแทน/agent, แผนกเคลม/ claim dept.,แผนกบริการทางโทรศัพท์/call center, แผนกลูกค้าสัมพันธ์/ customer service center, หน่วย/agency
	Proper noun	ปัญหาการเบิกค่าห้อง H&S/hospital benefit problem (H&S)	แม่ชีเทเรซ่า/Teresa nun (represent good behavior of people)	ตัวเหี้ย, ตัวเงินตัวทอง, แย้ (represent bad behavior of people)	ชื่อผลิตภัณฑ์/ 20TMAE, ค่าคุ้มครองรักษาพยาบาล/ H&S,ค่าคุ้มครองโรคร้ายแบบพิเศษ/ECIR	ข้อบังคับประกันภัย/ insurance regulation, กฎเรื่องประกันสะสมทรัพย์/Endowment Rules, สัญญาระยะยาว/ long term contract	เอไอเอ/AIA, ไทยประกัน/ Thaiprakan, เมืองไทยประกันชีวิต/Moung Thai life insurance
	qualitative/collective noun	ปัญหาทั้งหลาย/all of questions, คำถามพบบ่อย/ frequency quentions	-	-	กรมธรรม์ทั้งหมด/all of policies	ปัญหาการเคลมทั้งหลาย/all of claim problem	เหล่าตัวแทน/agent group
	abstract noun	-	-	-	-	-	-
	transitive/intransitive verb	ถาม/ask, สงสัย/wonder, อยากทราบ/eager to know, แนะนำ/advise, กรุณา/please	โปรด,พอใจ/love, ไว้วางใจ/trust	ผิดหวัง/ disappointed, โกรธ ,ฉุนเฉียว/cross	คุ้มครอง/protect, ถอดถอน/withdraw	สำรองจ่าย/provision, คุ้มครอง/protect	บริการ/service, สอบถาม/ enquiry
	adjective	-	ดี/good, เลิศ,ยอดเยี่ยม/excellent, great	ยอดแย่/worse, น่าเบื่อมากๆ/be tiresome, น่ารังเกียจ/disgusting	-	-	-
	adverb	-	อย่างมีสุข/well, อย่างดี/properly, nicely	อย่างรวดเร็วมากไป/ dramatically, ฉิบหาย/awfully, ถูลู่ถูกัง/forcefully, อย่างงุนงง/dazedly	-	-	-
Non-linguistic	Ambigous word, slang word	-	ชิวๆ/chill out, ติส/ bohemian style, แอ๊บแบ๊ว/to act childishly	สตรอเบอรี่/lie, ติงต๊อง/weird, ควาย/ stupid	-	-	-
Phrase	Phrase	ช่วยแนะนำแบบประกันสะสมทรัพย์/please recommend endowment life insurance product, can someone tell me?	มีใจบริการ/service mind, ขาดๆเกินๆ/ unbalance	หาหัวไม่เห็น/cannot meet, ไม่สามารถรับได้/cannot accept, รู้สึกแย่มากเหมือนกัน/feel bery bad, โกหกทั้งเพ/ absolutely lie	ต้องการซื้อแผนประกันคุ้มครองสุขภาพ20,25 ปี/wat to buy 20,25 years health protection plan	ปกปิดความจริงเรื่องสุขภาพ, แถลงสุขภาพไม่ตรงความจริง	ตัวแทนไม่มีใจบริการ/agent does not have service mind, ปิดบังข้อเท็จจริงของแบบประกัน/conceal the fact of plan

Table 5: Clue terms with linguistic and non-linguistic words example

4.3. Define sentiment scoring by SO-CAL concept

Identifying a polarity of sentiment word is useful for enhancing the power of the sentiment cube. The cube measurement declares that the only word count volume is not enough to show a sentiment severity. Making a score is more useful to get a clear viewpoint and show more analysis viewpoints. A scoring for sentiment words (lexicon & clue terms) in the additional dictionary was prepared by hand-tagging for giving a sentiment lexicon score by applying a scoring concept of Turney named SO-CAL [50]. Semantic Orientation CALculator (SO-CAL) obtains a process of word scoring from web-based search. The SO-CAL concept is used to identify the hand-ranked score using the same five scales for extremely positive and -5 scales for extremely negative, where 0 indicates a neutral word of noun, verb, and adverb. The scoring level expresses on adverb feeling group shows in SO value column in Table 6. The considerable thing of some parts of speech, such a noun or verb, is that it is possible to contain neutral and non-neutral connotations. This problem gains some benefits from the hand-tagging method. The calculation method of the intensifier modifier takes action like an amplifier increase the semantic intensity of words such as extraordinarily much stronger amplifier than rather.

For example, to define "*excellent"* has a SO value of *5*, *Conan's most excellent comic* would set an SO value equal: *5+ (5*100%) = 10*. In this work, the negation handling method is to make it ease by reverse a polarity of term in the lexicon, for example changing *good (+3)* into *not good (-3)*. All negation words are included, such as *not, none, nobody, never, nothing,* and other else.

Feeling Level	SO-Value
excruciatingly	-5
inexcusable	-3
foolishly	-2
satisfactorily	1
purposefully	2
hilariously	4

Intensifier	Modifier %
somewhat	-30%
pretty	-10%
really	+15%
very	+25%
extraordinarily	+50%
(the)most	+100%

Table 6: Sentiment scoring adjustment

When having determined the sentiment word in each sentence, score (s_i) can be computed as the sum of scores of the s_j is computed score (w_j) of all words $w_j \in s_i$ multiplied with their respective weights weight (w_j) under the rules of sentiment scoring adjustment:

$$Score\ (s_i) = \Sigma score\ (w_j) * weigh(w_j)$$

Example 1: SO-CAL sentiment calculation

Ex 1.1:

<Product aspect with negative sentiment>

*“**สิ่งที่แย่ที่สุด**คือ**ประกันตลอดชีพ** เพราะ**ใช้เวลานานมาก** กว่าจะได้รับทุนประกันคืน”*

*“**The worst thing** is **whole-life plan** because **it takes long time** to get sum assured return”*

=>*The worst = the most (bad: -5)* = -5+ (-5*100%) = -10

=>Sentiment score on whole-life plan = -10

=>Whole-life plan is clue terms of product aspect

Ex 1.2:

<Product topic with negative sentiment>

*“สิ่ง**ที่แย่ที่สุด**คือ**ประกันสะสมทรัพย์** โดยเฉพาะพวก**จ่ายเบี้ยหลาย ๆ ปี**”*

*“**The worst** thing is **endowment plan** especially, **long term payment**”*

=>*The worst = the most (bad: -5)* = -5+ (-5*100%) = -10

=>Sentiment score on endowment product = -10

=>Endowment plan is clue terms of product topic

<Claim aspect with negative sentiment>

“***สุดระอา*** *กับ**การเคลมประกัน**ของ**บริษัทA***”

”***Very fed up** with **claim process** of insurance **company A**”*

=>*Very = a lot (fed up*: -5) = -3+ (-5*100%) = -8

=>Sentiment score on claim assessment = -8

Remark: this sentence shows the relationship of claim aspect on company aspect in negative sentiment

Example 2: SO-CAL sentiment calculation + scoring of compositional semantic clue terms

Ex 2.1:

<Claim assessment aspect with compositional semantic clue-term negative sentiment>

*"ลูกค้ารู้สึก**ไม่พอใจอย่างมาก**ที่**บริษัทปฏิเสธการเคลมเงินประกัน** เพราะทางบริษัทได้พบว่ามี**การปกปิดโรคที่เป็นมาก่อนหน้า**"*

"Customer got ***very angry*** *that company* ***had rejected claim payment*** *because company had* ***found the concealment of the decease record****".*

=>very = a lot (angry: -5) = -3+ (-5*100%) = -8

=>Sentiment score on claim assessment = -8

=>"Rejected claim payment" and "concealment of the decease record" are claim assessment aspects. The compositional semantics clue terms scoring use in this case. Both phrases conduce to negative sentiment.

=>Total sentiment score = (-8) + score of "*reject payment"* + score of *"concealment of the decease record"*

Ex 2.2:

<Claim Problem sub-aspect with compositional semantic clue-term negative sentiment>

*"**บริษัทไม่จ่ายค่ารักษาทั้งหมด** เพราะทางบริษัทได้**สืบค้นประวัติเจอว่าคุณแม่เคยเป็นโรคมะเร็งมาก่อน**"*

"Company cannot pay all of medical fee *because company found* ***cancer health record*** *of mother".*

=>"*Company cannot pay all of medical fee"* is "claim problem" sub-aspect related to *"cancer health record"* claim aspect.

=> Sentiment score = score of "cannot pay all of medical fee" + 0

Remark:

Generally, *"cannot pay all of the medical fees*" does not have any negative feeling word; however, this phrase occurs in the claimed domain. It shows negative sentiment. In this case, compositional semantic clue-term, which was tagged from Sansarn tagging tools, is able to support this case. So this phrase show negative sentiment in the 'Claim' aspect and score equal to SO-CAL assignment. "*cancer health record*" does not have any feeling score; however, this is a keyword to show the relationship of claim problem with cancer decease. Finally, of this step of information extraction, corpus and lexicon are already prepared to extract features in each aspect from each chunk of a sentence and then determine sentiment degree on them.

5. Exploiting Linguistic Knowledge for Sentiment Classification

This part makes an experiment to exploit linguistic knowledge. It aims to improve topic and sentiment classification in our main corpus for automatically assigning appropriate labels on an appropriate class by training the classification models and language resources using our design clue terms such as linguistic clue term etc., to extract sentiment from texts. We used some standards supervised automatic text classification methods such as Naïve Bayes, Decision Tree (J48), and Sequential minimal optimization (SMO) to evaluate the performance. The results from our approach have some improvement in the accuracy rate.

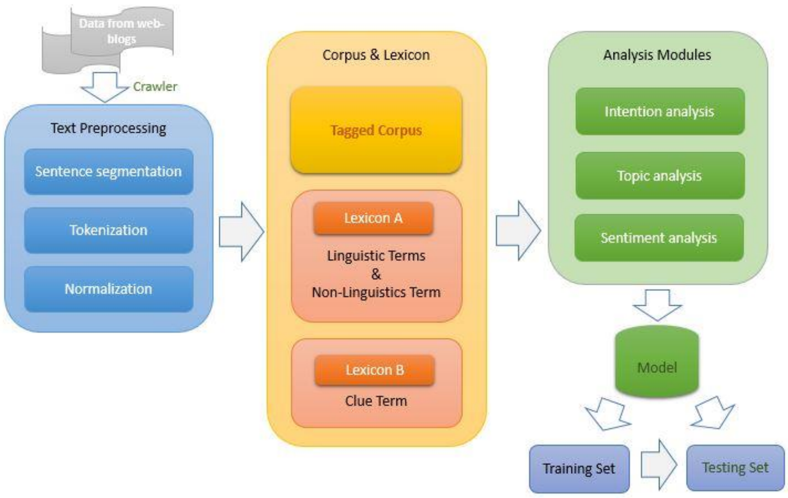


Figure 23: The example of collection of clue terms for intention and topic analysis

As corpus preparation above section, the corpus from web-blogs consists of linguistics terms and non-linguistic terms, including Thai synonymous words, transliterated words, abbreviations, and slang. Corpus was prepared by divided sentences are separated from each users' message in each topic using space and determined length of Thai characters per line around 200-300 characters. The number of divided testing sentences is 4,668. The tagged corpus was prepared from random discussion topics and was annotated into three aspects of intention, topic, and sentiment. After that, we do the tokenization and normalization process and then extract key feature terms of phrases from a given text. (Refer to Figure 23), Lexicon preparation in sentiment classification. The design and completeness of the lexicon base affect the performance of the sentiment analysis. Therefore, the corpus preparation process is a necessary process to make performance more reliable [7]. We stripped out HTML tags from the corpus and made a part in the chunks of words. The tagged corpus was annotated in three aspects, intention, topic, and sentiment, to support main analysis components under our method; lexicons are provided as follows.

The volume of tagging corpus in topic analysis modules (service, product, and regulation) is 46.36%, 7.72%, and 39.48%, respectively. The main lexicon is LEXiTRON. It is a general word from the Thai dictionary that has 35,933 words. Particular lexicon volume from this approach (special clue terms and general clue terms) is 2,439 terms. Sentiment lexicon volume is 1,434 terms.

We experiment with our three kinds of the lexicon, including with Lexicon semantics clue Terms, Complement clue Terms, and Compositional semantics clue terms. We demonstrate an automatically assign appropriate labels on an appropriate class by training the classification models and language resources manually tagged labels, including special clue terms construction. Besides, we try to provide more sentiment analysis by giving a score on special clue terms to indicate the severity level of each customer topic from a cross-relationship of relevant sentences.

Table 7 shows the analysis results of the relationship between main topics in corpus and lexicon mapping by special clue terms. The hidden relationships in each topic are revealed in the percentage of the number of mapping words from special clue terms. The strength of the relationship indicates with high cross percentage. That means they have a high relationship with customers' subjectivity. Each row in *table 7* is corpus details. There are claim, service, regulation, product, agent, and intention. All of them are criticized topics in web-blogs. Each column in *table 7* is mapping results from all of the lexicons from our approach. The results show the percentage of each lexicon mapping in each corpus. One corpus can find word mapping with lexicon more than one in each divided sentence. Some divided sentences have sentiment words, but some divided sentences cannot find.

For example of the results from table 7:

Corpus	Topic Analysis							Intention Analysis		Sentiment Analysis			
	Service (%)				Prod (%)	Regulation (%)		Ques(%)	Inf (%)	Like (Score points)	Like (%)	Not Like (Score points)	Not Like (%)
	claim	agent	CSC	CCC		REG	comp						
Claim	**19.30**	5.26	0.00	0.00	7.02	**40.35**	3.51	**8.77**	3.51	96	10.53	-18.00	1.75
Service	**27.61**	**22.70**	0.00	0.61	3.07	**19.63**	4.91	6.13	0.00	162	7.36	-276.00	7.98
Regulation	**14.00**	2.39	3.59	1.08	6.34	**32.42**	**7.42**	4.19	4.43	1,920	13.76	-1482.00	10.41
Product	**16.67**	0.00	0.00	0.00	**16.67**	**27.78**	2.78	**13.89**	0.00	114	22.22	0.00	0.00
Agent	0.00	**30.73**	0.00	0.46	**13.30**	**17.89**	5.50	6.88	0.00	636	17.89	-252.00	7.34
Intention	5.79	**23.14**	0.00	0.00	0.00	**21.00**	1.65	**10.74**	0.83	420	19.01	-396.00	17.36

Table 7: The analysis results in relation in topics, intentions and sentiments

Case A: Agent corpus can find the word mapping in agent lexicon (itself) 30.73%. An agent has a high impact on company regulation, 17.89%, and product 13.3%.

Case B: Claim corpus can find the word mapping itself 19.30%. A claim has a high relationship on company regulation as 40.35%.

Case C: Product corpus can find the word mapping itself 16.67%, and people talk about the claim process in product corpus 16.67% and talk about company regulation in product 27.78%. And the results show people ask questions in product corpus 13.89%.

A high percentage of strong relationships should be the first issue is needed insurance company focuses on. In this analysis, we found that the first important issue is that company regulation impacts insurance customer activities. The second is the agent's service to the customer, and the third is a claim issue. For sentiment scores in the like and not like column shows the problem on service topic and high volume is related to company regulation issues. One of the reasons may come from the customer knowledge in insurance is not enough, so providing enough life insurance knowledge may increase customer satisfaction.

The vital thing of in-depth analysis is the annotated special clue lexicons design. Suppose they provide a sub-level of each issue. In that case, the analysis results might reveal each problem in detail, so more study in business knowledge is needed to design special clue terms. To perform an evaluation task, we considered three standard supervised algorithms: Naïve Bayes, Decision Tree (J48), and SMO. The performance of each three methods is compared between baseline results (general) and our methodology results (general + clue) inaccuracy rate evaluation as *in Table 8*. Three algorithms are performed by using 10-fold cross-validations.

Topic	Algorithm	Term feature	Accuracy (%)
Product	Naïve Bayes	General	93.96
		General + Clue	93.83
	Decision Tree (J48)	General	92.29
		General + Clue	92.29
	SMO	General	94.34
		General + Clue	94.60
Regulation	Naïve Bayes	General	68.77
		General + Clue	69.41
	Decision Tree (J48)	General	68.12
		General + Clue	69.28
	SMO	General	70.95
		General + Clue	73.52
Customer Service	Naïve Bayes	General	75.71
		General + Clue	75.06
	Decision Tree (J48)	General	75.84
		General + Clue	76.22
	SMO	General	75.45
		General + Clue	77.12

Table 8: Accuracy rate of each algorithm

We found some improvement of accuracy rate in almost algorithms, especially SMO algorithms and except on product topic and customer service topic in Naïve Bayes algorithm.

6. Information extraction process from corpus and lexicon

Breaking down opinions into aspect level:

We extracted words according to identified aspects or sub-aspects under hierarchical form based on lexicons and clue terms to generate insightful meaning. The corpus that was loaded from the websites is unstructured data with the narrative format, divided into a chunk of sentences, as stated in the previous section.

Figure 24 shows the extraction of words inside a sentence that matching with the lexicon. Start from blog/document level:d split to be sentence level:s and then divided in matching word/term:t. One chunk of the sentence has some words which were matched with the lexicon. That is, sub-aspects under some aspects under a hierarchical structure. The output of clue terms of aspects/sub-aspects mapping on a chunk of sentences shows that the adjacent aspects/sub-aspects have the meaning associated with each other.

The results of this step, tracking word matched on the lexicon, we get aspect/sub-aspect/ (multi-dimension), hierarchical structure (sub-aspect), multi-measurements as stated in the previous section. The number of word counts and sentiment scores were calculated as SO-CAL formula to present the opinion results breaking down in each aspect and sub-aspect. The summarization of the score is a group under the same sub-aspect. We extracted words according to identified aspects or sub-aspects under hierarchical form based on lexicons and clue terms. The corpus that was loaded from the websites is unstructured data with the narrative format divided into a chunk of sentences, as stated in the previous section.

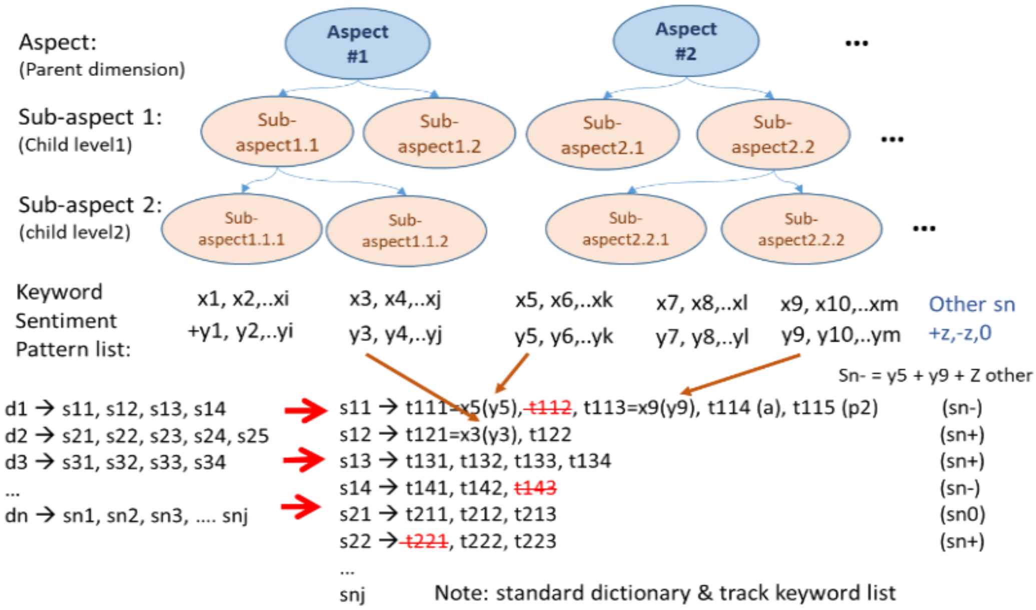


Figure 24: Process of aspect/sub-aspects mapping using lexitron dictionary and clue terms

Remark: *d is document, s is sentence, t is term, x is clue term, y is score of each clue term, sn = sentiment score of each chunk*

There is a notice of unstructured data tracking with clue terms. The aspect tracking is unable to match entirely for every aspect on a chunk of a sentence because it depends on the sentence length limitation. Therefore, it has data incompleteness. A null value was found in the unmatched field on the fact table. However, unmatched fields remain some associations between the other aspects, so the analysis viewpoint should not be ignored. On the contrary, we can match multiple clue terms of the same aspect in the same chunk. It shows that this chunk gives precedence to these aspects more than the others. Especially, the cube property provides an analysis viewpoint across a different layer of the hierarchy, so null values of sub-aspect (child level) are able to analyze in aspect (parent level) for finding a related relationship with the other dimensions.

7. Design of the Fact tables population

Refer to Figure 17, after we have finished in part of unstructured data preparation by performing information extraction under aspect-based sentiment analysis. Then data preprocessing on OLAP is required for the data population process. The general process of data preprocessing such as data cleansing, extract transform, and load (ETL) is necessary to prepare structured data to be more pattern as analysis purpose for a populated table in the database called transaction tables. Data was loaded into transaction tables are the results from the information extraction process. Fact tables are generally loaded from transaction tables such as order tables or transactional files. Therefore the number of rows to update or insert in a load is much larger than in dimensions. The core of loading fact tables is to change the natural keys into surrogate keys [89]. To re-arrange data patterns and calculated measurements followed in analysis concepts are required structured query language (SQL) and ETL to populated fact tables. The different viewpoints can transfer from transaction tables to fact tables from a different point of view by using SQL scripts and operations. The different design of fact tables by SQL makes the data value of the fact table more meaningful in different ways of analysis. These fact tables are used for building multidimensional sentiment cube. A data cube is a multidimensional structure of a data visualization that is used in analytical data warehouses using the Online Analytical Process (OLAP). The design of the fact table directly impacts cube model characteristics, so cube performance and analysis views depend on population table design. Those of fact tables designate for taking the place of a multidimensional sentiment conceptual model as below.

define lifeINS_cube [*company, product, service_touchpoint, service_key_performance, claim_type, claim_problem, impact_level*]:
service_performance_count = sum(*service_word_count*),
service_performance_score = sum(*service_word_score*),
claim_problem_count=sum(*claim_problem_word_count*),
claim_problem_score= sum(*claim_problem_word_score*),
impact_level_count = sum(*impact_level_count*),
impact_level_score = sum(*impact_level_score*),
negative_count = sum(*negative_count*),
negative_score = sum(*negative_score*),
positive_count = sum(*positive_count*),
positive_score = sum(*positive_score*),
sentence_count = sum(*sentence_count*)
define dimension *company* as (*company_name*),
define dimension *product* as (*Life, Rider, Affinity, special, other*)
define dimension *service_touchpoint* as (*internal-_service, external_service*)
define dimension *service_evaluation* as (*reliability, consistency, response, knowledge, characteristic*)
define dimension *claim_type* as (*minor_claim, major_-claim*)
define dimension *claim_problem* as (*advance_money, reject_payment, conceal, fax, long_service, misunder-stand_coverage,misunderstanding_in_protection, misunderstanding_in_period, misunderstand_in_exception*)
define dimension *impact_level* as (*make_sue_on_court, make_petition_for_OIC, make_accusation_agent, loan,invalid_policy, void_policy, surrender_policy, policy_reject, policy_cancellation*)

In this paper, we performed three fact table design patterns to find out the good points and weak points from our design examples. Master table is one kind of table that looks up as a hierarchical structure. This kind of table is used for joining with transaction table or fact table for analysis. In our case, we designed a master table that refers to our analysis viewpoint, which relates to aspects (parent level) and sub-aspects (child level). Our master tables compose the table Product, Company, Service Touchpoint, Service Evaluation Factors Process and Operation, and Dissatisfaction Impact Type as Table 4.

Step 3. Utilize and Analysis Results

This step designs three types of a cube as following to utilize a multi-dimensional sentiment cube to analyze and find knowledge.
Three types are

1. Single aspect analysis (Pattern A)
2. Single Aspect Co-occurrence analysis (Pattern B)
3. Multiple Aspect Co-occurrence analysis (Pattern C)

3.4.1 Single aspect analysis (Pattern A):

id	doc_id	own_id	st_id	sap_cd_i	sap_cnt_i	sap_sc_i
1	01	01	01	cd_{11}	cnt_{11}	sc_{11}
2	01	01	02	$...cd_{1m}$	$...cnt_{1m}$	$...sc_{1m}$
3	01	02	01	cd_{21}	cnt_{21}	sc_{21}
4	01	02	02	$...cd_{2n}$	$...cnt_{2n}$	$...sc_{2n}$
5	02	01	01	$...cd_{ij}$	$...cnt_{ij}$	$...sc_{ij}$
6	02	01	02	$ng\_cd_{11...ij}$	$ng\_cnt_{11...ij}$	$ng\_sc_{11..ij}$
...j	02	02	03	$ps\_cd_{11..ij}$	$ps\_cnt_{11..ij}$	$ps\_sc_{11..ij}$

Table 9: An example of fact table aspectPattern_A

This pattern is to analyze aspects by extract a single sub-aspect. The fact table's key is document id, owner id, sentence id, and sub-aspect code. Unique key of this table is doc_id||own_id||st_id||sap_cd. In general, one sentence is possible to find matching words in some aspects or sub-aspects. If this pattern finds more than one sub-aspect in the same sentence, the SQL script will generate it by inserting a new row of fact tables. It means one row of the table has only one sub-aspect to be an element of a concatenated key. That's why a unique key is required sub-aspects to be one key. The primary dimension for analysis is a single sub-aspect column that links with the master table of aspects. As a hierarchy of master table, the analysis viewpoint will follow the lookup step on hierarchical from child to parent. Measurements are the number of words count of each sub-aspect and score of terms count of each sub-aspect. So the value of multi-measurements can roll-up as a hierarchy of master table.

Example of this case:
"<u>อยากยกเลิกกรมธรรม์</u>ของ<u>บริษัทA</u>มาก ทำเรื่อง<u>เคลมให้ก็ช้า</u> <u>ตัวแทนประกันไม่ช่วยเหลือ</u> ติดต่อหน่วยประกันก็ไม่ได้ สุดท้ายยัง<u>โดนปฏิเสธการเคลม</u>อีก"
"I <u>would like to *make policy cancellation*</u> of the <u>company A</u>, make <u>*claim process* very slow</u>. <u>*Agent* did not help</u>. <u>*Agency cannot contact*</u>. Finally, <u>claim was *rejected payment*</u>."

"Make policy cancellation" is Medium Impact Type sub-aspect under Dissatisfaction Impact Type aspect *(code 'PC', phrase score -5).*
"company A" is Company aspect *(code 'comA', word score = 0)*
"agent did not help" is Consistency sub-aspect under Service Evaluation aspect *(code, 'CONS', phrase score = -3)*

"Agency cannot contact" is Response sub-aspect under Service Evaluation aspect *(code ''RESP', phrase score = -3)*
"claim process very slow" (code 'CLM', phrase score = -4) is General Claim > Claim Assessment under Process and Operation aspect
"rejected payment" (code 'RJ', phrase score = -5) is Reject Payment sub-aspect>Claim Problem under Process and Operation aspect

Refer to SQL statement as below.

```
CREATE fact table aspectPattern_A as
SELECT key_id11…1j, (id, doc_id, own_id, st_id)
CASE WHEN count of sub_aspect11..1m <> 0
THEN  'sub_aspect11..1m_cd'                    AS    sap_cdi,
          sub_aspect11..1m_count                AS    sap_cnti,
          sub_aspect11..1m_score                AS    sap_sci,
UNION
CASE WHEN count of sub_aspect21..2n <> 0 THEN 'sub_aspect21..2n_cd'
                                                AS    sap_cdi,
          sub_aspect21..2n_count                AS    sap_cnti,
          sub_aspect21..2n_score                AS    sap_sci,
UNION
CASE WHEN count of negative11..ij <> 0 THEN 'negative11..ij_cd'
                                                AS    sap_cdi,
          negative11…1i_count                   AS    sap_cnti,
          negative11…1i_score                   AS    sap_sci,
UNION
CASE WHEN count of positive11..ij <> 0 THEN 'positive11..ij_cd'
                                                AS    sap_cdi,
          positive11…1i_count                   AS    sap_cnti,
          positive11…1i_score                   AS    sap_sci,
Repeat CASE of sub_aspect_id = 3…i
FROM populate table
```

3.4.2 Single Aspect Co-occurrence analysis (Pattern B):

id	doc_id	own_id	st_id	sap_cd_1	sap_cnt_1	sap_sc_1	$...sap\_cd_i$	$...sap\_cnt_i$	$...sap\_sc_i$	ng_cnt	ng_sc	ps_cnt	ps_sc
1	01	01	01	cd_{11}	cnt_{11}	sc_{11}	0	0	0	ngc_1	ngs_1	psc_1	pss_1
2	01	01	02	0	0	0	$...cd_{i2}$	$...cnt_{i2}$	$...sc_{i2}$	ngc_2	ngs_2	psc_2	pss_2
3	01	02	01	cd_{13}	cnt_{13}	sc_{13}	0	0	0	ngc_3	ngs_3	psc_3	pss_3
4	01	02	02	0	0	0	$...cd_{i4}$	$...cnt_{i4}$	$...sc_{i4}$	ngc_4	ngs_4	psc_4	pss_4
...j	01	02	02	cd_{1j}	cnt_{1j}	$sc1_{1j}$	$...cd_{ij}$	$...cnt_{ij}$	$...sc_{ij}$	ngc_i	ngs_i	psc_i	pss_i

Table 10: An example of fact table aspectPattern_B

This pattern analyzes aspects by extracting a single sub-aspect with remaining the other sub-aspects' structure in the same record with value zero. The fact table's key is document id, owner id, sentence id, and sub-aspect code. Unique key of this table is doc_id||own_id||st_id||sap_cd. We used matching words between each sentence and lexicon to identify aspects/sub-aspects, which is a column (dimension) in a fact table.

The purpose of pattern B is to keep relation with other sub-aspects in the same record. Even if the other records have some zero value in some records, the column (aspect) still can look up as the hierarchy of the master table with some value in that column. Several multi-dimensions are equal to some analysis sub-aspects. The number of multi-measurements is similar to all of the analysis sub-aspects, which contain the group of many word counts of each sub-aspect and the score of each sub-aspect. Refer to SQL statement as below.

CREATE fact table aspectPattern_B as
SELECT key_id$_{1\ldots j}$, *(id,doc_id,own_id,st_id)*
CASE WHEN count of sub_aspect$_1$ <> 0 AND count of sub_aspect$_{2\ldots i}$ = 0
THEN ‘sub_aspect$_1$_cd’ AS sap_cd$_1$,
sub_aspect$_1$_count AS sap_cnt$_1$,
sub_aspect$_1$_score AS sap_sc$_1$,
negative$_1$_count AS ng_cnt1,
negative$_1$_score AS ng_sc1,
positive$_1$_count AS ps_cnt1,
positive$_1$_score AS ps_sc1,
$0_2,\ldots,0_i$ AS sap_cd$_{2\ldots i,}$
$0_2,\ldots,0_i$ AS sap_cnt$_{2\ldots i,}$
$0_2,\ldots,0_i$ AS sap_sc$_{2\ldots i}$
UNION
SELECT key_id$_{j+1\ldots m}$, *(doc_id,own_id,st_id)*, ‘’ AS sap_cd$_{1,}$ 0 AS sap_cnt$_{1,}$ 0 AS sap_sc$_{1,}$
CASE WHEN count of sub_aspect$_2$ <> 0 AND count of sub_aspect$_1$ = 0 AND count of sub_aspect$_{3\ldots i}$ = 0 THEN ‘sub_aspect$_2$_cd’ AS sap_cd$_2$,
$0_3,\ldots,0_i$ AS sap_cd$_{3\ldots i,}$
sub_aspect$_2$_count AS sap_cnt$_2$,
sub_aspect$_2$_score AS sap_sc$_2$,
negative$_2$_count AS ng_cnt$_2$,
negative$_2$_score AS ng_sc$_2$,
positive$_2$_count AS ps_cnt$_2$,
positive$_2$_score AS ps_sc$_2$,
$0_3,\ldots,0_i$ AS sap_cd$_{3\ldots i,}$
$0_3,\ldots,0_i$ AS sap_cnt$_{3\ldots i,}$
$0_3,\ldots,0_i$ AS sap_sc$_{3\ldots i}$
Repeat CASE of sub_aspect_id = 3…i
FROM populate table

3.4.3 Multiple Aspect Co-occurrence analysis (Pattern C):

id	doc_id	own_id	st_id	sap_cd$_1$	sap_cnt$_1$	sap_sc$_1$	…sap_cd$_i$	…sap_cnt$_i$	…sap_sc$_i$	ng_cnt	ng_sc	ps_cnt	ps_sc
1	01	01	01	cd$_{111}$||cd$_{112}$	cnt$_{111+112}$	sc$_{111+112}$	…cd$_{i1}$	…cnt$_{i1}$	…sc$_{i1}$	ngc$_1$	ngs$_1$	psc$_1$	pss$_1$
2	01	01	02	cd$_{12}$	cnt$_{12}$	sc$_{12}$	…cd$_{i2}$	…cnt$_{i2}$	…sc$_{i2}$	ngc$_2$	ngs$_2$	psc$_2$	pss$_2$
3	01	02	01	cd$_{13}$	cnt$_{13}$	sc$_{13}$	…cd$_{i31}$||cd$_{i32}$	…cnt$_{i31+i32}$	…sc$_{i31+i32}$	ngc$_3$	ngs$_3$	psc$_3$	pss$_3$
4	01	02	02	cd$_{14}$	cnt$_{14}$	sc$_{14}$	…cd$_{i3}$	…cnt$_{i4}$	…sc$_{i4}$	ngc$_4$	ngs$_4$	psc$_4$	pss$_4$
…j	01	02	02	cd$_{1j\,1}$||cd$_{2jK}$	cnt$_{1j\,1+\ldots+2jK}$	sc1$_{j1}$ $_{sc}$1$_{j3}$	…cd$_{ij}$	…cnt$_{ij}$	…sc$_{ij}$	ngc$_i$	ngs$_i$	psc$_i$	pss$_i$

Table 11: An example of fact table aspectPattern_C

This pattern is to analyze aspects by extract multiple sub-aspects in the same records. In tracking multiple-sub-aspects in the same records, sub-aspect code (column_name) will be represented with new concatenated code. The key of fact table consists of document id, owner id, and sentence id. The unique key of this table is doc_id||own_id||st_id. In this case, sub-aspect will not be a key because one record is represented to get value from one sentence.

Refer to SQL statement as below.

CREATE fact table aspectPattern_C as
SELECT key_id$_{11\ldots1j}$, *(id, doc_id, own_id, st_id)*
[[CASE WHEN count of sub_aspect$_{11}$ <> 0 AND count of sub_aspect$_{12\ldots1i}$ = 0 THEN
'sub_aspect$_{11}$_cd' AS sap_cd$_1$,
sub_aspect$_{11}$_count As sap_cnt$_1$,
sub_aspect$_{11}$_score AS sap_sc$_1$],
[CASE WHEN count of sub_aspect$_{11}$ <> 0 and count of sub_aspect$_{12..1n}$ <> 0 THEN
'sub_aspect$_{11}$_cd'||'sub_aspect$_{12\ldots1n}$_cd' AS sap_cd1,
SUM(sub_aspect$_{11\ldots1i}$_count) As sap_cnt$_1$,
SUM(sub_aspect$_{11\ldots1i}$_score) AS sap_sc$_1$,],]
SUM(negative$_{11\ldots1i}$_count) AS ng_cnt1,
SUM(negative$_{11\ldots1i}$_score) AS ng_sc1,
SUM(positive$_{11\ldots1i}$_count) AS ps_cnt1,
SUM(positive$_{11\ldots1i}$_score) AS ps_sc1,
Repeat CASE of sub_aspect_id = 2…i
FROM populate table

Remark:
sap_cd$_{1..i}$ is sub-aspect code$_{1..i}$, sap_cnt$_{1..i}$ is sub-aspect word count$_{1..i}$, sap_sc$_{1..i}$ is sub-aspect score$_{1..i}$, ng_cnt$_{1..i}$ is no.of negative words, ng_sc$_{1..i}$ is score of negative words, ps_cnt$_{1..i}$ is no.of positive words, ps_sc$_{1..i}$ is score of positive words,
id is record id, doc_id is post/document id, own_id is post woner id,
st_id is sentence_id; j is record_id ; i is aspect_id

Step 4. Sharing

There are many ways of making knowledge sharing. For example, it may prepare by document, database, or technology in part of explicit knowledge. For tacit knowledge, co-team setting across co-workers, sharing activities and innovation, learning community, and job position changes can be utilized for the knowledge sharing process. However, in this study, we use the method of process sharing using a questionnaire. There are many previous works to use knowledge sharing by conduct the questionnaire. Several Thai scholars have examined customer satisfaction with life insurance [66], [74], [94]. However, some gaps remain in the current research, which customer satisfaction evaluation, using outside-in knowledge, could partially bridge. For instance, even when companies implement customer satisfaction improvement strategies, customer complaints expressing dissatisfaction remain on social media.

Therefore, this research examines the dissatisfaction with Thai life insurance services and analyzes the severity level of dissatisfaction for each problem. It also identifies problem-solving methods that customers appreciate. In-depth analysis for focus on problem-solving is necessary for the service issues that dissatisfied customers express on social media. This study analyzes the severity of customer satisfaction with service dimensions based on negative sentiments on social media using a questionnaire survey. To the end, it deploys the analytical framework from Figure 25.

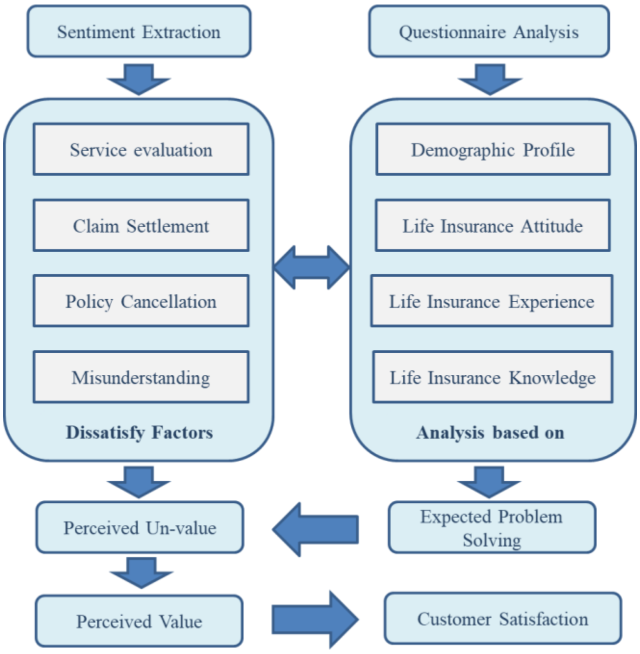


Figure 25: Customer Sentiment Analytical framework

We prepared a questionnaire by first tracking the dissatisfaction issues extracted from negative customer sentiments on social media based on this framework. The "*Sentiment extraction tool*" was created for tracking negative and positive opinions on web-blogs on Thai social media site "Pantip.com." The tools used factor keywords related to life insurance service in the tracking process [61]. Using the results of sentiment extraction, we prepared questions on these dissatisfaction issues. The results identified four dissatisfaction issues: service quality, claim settlements, policy cancellations, and misunderstandings. All these issues devalue the service in the customers' minds.

In the next step, we analyzed the questionnaires, considering the four dissatisfaction issues across customer characteristics such as demographics, life insurance attitude, life insurance experience, and life insurance knowledge. Besides, this survey included questions about problem-solving methods, which respondents rated the highest. The implication is that such methods as agreed to by most customers might enable efficient problem-solving based on customer opinions. Using customer engagement to identify ideas for solving problems is another SCRM concept. After implementing the problem-solving methods, a customer will shift from an un-valued service perception to a valued one. Ultimately, customer satisfaction could increase. This study conducted the questionnaire based on sentiment extractions tools and multi-dimensional sentiment cube analysis results in four main parts, which consists of service evaluation factors, claim settlement problems, policy cancellation motive, and misunderstanding issues between customer and service provider.

- **Service evaluation factors** are the service factor dimensions perceived and judged by customers when receiving such services.
- **Claim settlement** is the insurer's obligation in terms of an insurance contract with the policyholder (insured) to pay compensation. When the insured's policy covers the claim, a predetermined amount of money is claimed; namely, the sum assured (S.A.) under the claim assessment principle [44].

  - There are three types of claims in life insurance policies
    - Survival Benefit Claim such as Out-Patient Department – OPD, In-Patient Department – IPD, Hospital Benefit –HB, Critical Illness – CI, Cancer Benefit – CB, Total Disability, and so on.
    - Maturity Benefit Claim
    - Death Benefit Claim

- **Policy cancellation** is the end of an inforce policy status before the expiration of the policy. In this case, we review policy cancellation and policy lapses due to customer requests.
  - *Policy cancellation* can be done through a request from the policyholder (owner/insurer) or the life insurance company. In some cases, the policyholder requests policy cancellation due to dissatisfaction with the product or service. The life insurance company may also cancel the policy if it finds that the policyholder has failed to comply with the contract; for example, a customer conceals health records.
  - *Policy lapsation* occurs automatically when the policyholder discontinues a premium payment, so that all benefits of the policy also stop automatically.
  - However, the termination of the policy has a negative effect on both the policyholder and the life insurance company. The policyholder loses the insurance policy's protection, and the accumulated cash value is also lost; the life insurance company even loses the money due; thus, termination is not profitable for either party.

- **Misunderstandings of life insurance principles and regulations** between policyholders and service providers include controversial problems that generate lawsuits or cancellations of life insurance policies, as expressed on social media.

The objectives of the **questionnaires conduct** are as follows:

1. To measure customer satisfaction with the service evaluation factors to prioritize the most important problems.
2. To study the causes of policyholder dissatisfaction with claim settlements, policy cancellation motives, and reasons leading to misunderstandings with the life insurance principles and regulations that generates dissatisfaction relevant to users' negative sentiments on Thai social media.
3. To study significant factors, such as demographic characteristics, life insurance experience, knowledge, and attitudes that influence service evaluation factors.
4. To study the appropriate problem-solving methods for dealing with dissatisfaction.

Step 5. Capture and Learning

To capture knowledge sharing between customer sentiment and respondents who replied to the questionnaire, the results of customer sentiment knowledge sharing got from work experience, the records of the knowledge apply from knowledge sharing or knowledge warehouse, etc. To study customer sentiment related to our main target, we evaluated the questionnaire results under four dimensions of the respondents' qualifications: demographic characteristics, life insurance experience, life insurance knowledge, and life insurance attitudes of the Thai respondents.

Demographic characteristics:
The demographic characteristic consisted of gender, age, education level, occupation, and income. These were important indicators that help determine the target group for analysis.

Life insurance experience:
The experience gained from observing, learning and remembering is considered a life lesson. Each person has a different experience. However, many people can share an experience together. Learning by experience can become knowledge (20th Century Chambers dictionary). This study compared respondents who had work experience in life insurance, such as staff or agents, including those taking a particular class on life insurance, with respondents who did not have any experience with life insurance.

Life insurance knowledge:

Knowledge is an integration of information based on values, experiences, contextual information, and expert understanding, providing a framework for evaluating and consolidating novel information and experiences which creates in the individual minds. [23]. This study compared the results of respondents who had at least basic knowledge of life insurance with respondents who did not have such basic knowledge. We requested that respondents evaluate their own level of understanding, which we tested through basic questions on their understanding of topics such as premium, sum assurance, and product (plan) type.

Life insurance attitude:
Attitude is a feeling or opinion about something or toward someone or an expressive behavior caused by this feeling or opinion. The "attitude-toward-object model" presents consumer attitudes toward products or brands by evaluating specific product properties' reliability. When a customer is satisfied with the product or service, he/she believes it has sufficient qualifications. Therefore, customers will positively evaluate the object; however, if they have a negative attitude and are not satisfied with the product or service. They will negatively evaluate the product or service (Muhammad et al., 2015) [53]. This study compared the responses among respondents who had positive, neutral, or negative attitudes toward life insurance.

Research Method: Customer Sentiment Evaluation Questionnaire
A descriptive research design was adopted to analyze customer satisfaction with Thai life insurance. This study is based on primary data collected using a questionnaire

spread via the Google document platform. We created a standard adaptive questionnaire and separated it into parts A, B, C, D, and E. Part A contained the demographic characteristics of respondents. Part B contained the life insurance attitude of the respondents. Part C comprised the basic life insurance knowledge of the respondents. Part D consisted of the service factors, claims service, policy cancellations, and misunderstandings. Finally, part E contained an evaluation of service improvement methods. The sample size was 374, and suitable statistical tools such as *percentage analysis, average mean score values, and ANOVA* at a significance of 0.05 were adopted. The collected data were analyzed, and the results were interpreted as follows.

Data Collection Method

This study's data were acquired from 374 Thai respondents with different demographic characteristics (gender, age, education, occupation, and income). The respondents had different life insurance experiences, knowledge of life insurance, and attitudes toward life insurance. Our questionnaire's target respondents were customers, the insured (policy owners), prospective customers, or non-customers. Table 12 presents the respondent information.

		Frequency	**Percent**
	Female	209	55.9
Gender	Male	165	44.1
	Total	**374**	**100.0**
	20 yrs or less	2	0.5
	21 - 30 yrs	31	8.3
	31 - 40 yrs	86	23.0
Age	41 - 50 yrs	149	39.8
	51 - 60 yrs	91	24.7
	61 yrs or up	14	3.7
	Total	**374**	**100.0**
	Less than Bachelor's Degree	23	6.1
	Bachelor's Degree	230	61.5
Education	Master's Degree	111	29.7
	Doctorate	10	2.7
	Total	**374**	**100.0**
	Self-employed	53	14.2
Occupation	Student	8	2.1
	Employee	259	69.3
	Government Official	26	7.0
	Retired person/ Unemployed	28	7.5
	Total	**374**	**100.0**

		Frequency	**Percent**
	No income	11	2.9
	25,000 or less	48	12.8
	25,001 - 50,000	139	37.2
Income (BTH)	50,001 - 100,000	99	26.5
	100,001 - 200,000	39	10.4
	200,001 or up	38	10.2
	Total	**374**	**100.0**
Experience	Have	167	44.7
	Do not have	207	55.3
Experience Knowledge level	**Total**	**374**	**100.0**
	Less	69	18.4
	Medium	167	44.7
Knowledge level Attitude	High	138	36.9
	Total	**374**	**100.0**
	Positive	321	85.8
Attitude	Neutral	38	10.2
	Negative	15	4.0
	Total	**374**	**100.0**
	Total	**374**	**100.0**

Table 12: Respondent characteristics

Step 6. Create and Leverage

There is a new **knowledge asset** from the **knowledge exchange process** using technology support from all of the processes as state above. This knowledge utilizes to help make a decision process of problem-solving method or drives innovation of work processes from knowledge creation. We are applying learning knowledge and new experiences to continue to circulate continuously.

Exploiting customer sentiment relate to customer knowledge is one kind of new knowledge creation. Customer sentiment-related knowledge involved in the interactions between a company and customers on a social media base can be classified into three types.

1. **Knowledge about customer sentiment** – is a kind of knowledge that a company attains to know the customer's opinion on a product or service, including the customer's profile who expresses that feeling. However, such kind of this information, when extracting customer sentiment from social media, the customer profile is anonymous. The knowledge sharing using questionnaires can fulfill the customer profile gap.

2. **Knowledge from customer sentiment** is the knowledge feedback and their contributions idea by customers. We can realize what customers feel about products or services from expressed words on social media from the sentiment extraction tool and sentiment cube. It also reveals the hidden problems and requires solving the questions very fast if they got very negative sentiment. The questionnaire is related to those problems we can know why and what customers think about those problems.

3. **Knowledge for customer sentiment** is the knowledge a company purposively provides to customers or the knowledge shared among customers, increasing customer satisfaction on product or service.

Step 7. Implementation

As the word "The great end of knowledge is not knowledge but action" (DOPA KM Team, Mason & Mitroff, (1973)) Knowledge dissemination (an implementation) is needed to achieve the goal which has organized knowledge management activities (KM Process) and change management process activities simultaneously, with the expectation that this knowledge management plan will be an important starting point for the implementation in the scope of Target knowledge management in other matters and lead to a sustainable learning organization.

Step 8. Learning and Monitoring

Monitoring (Plan the next iteration) is an activity that looks forward to testing and evaluating each round's objectives. To follow up on new knowledge creation results to prove this knowledge is sustainable usage and finding new ways to create new knowledge.

Chapter 4
Experiment Results

The results for answer the research questions are the followings

Main Research Question: How to utilize the knowledge on customer sentiment for improving customer satisfaction in Thai life insurance?
Sub Research Question 1: How to develop analysis tool for extracting customer sentiment knowledge from social media?
Sub Research Question 2: What are the issues relate to customer sentiment in Thai life insurance from knowledge sharing process?
Sub Research Question 3: What kinds of knowledge on customer sentiment have been utilize for knowledge sharing process between service provider and customer?

Main Research Question: How to utilize the knowledge on customer sentiment for improving customer satisfaction in Thai life insurance?

Our knowledge management methodology utilizes the knowledge on customer sentiment by defining the customer sentiment which obtained from the Outside-In and Inside-Out management approach to reveal hidden problems of Thai life insurance service and to get knowledge for finding out the appropriated methods to increase customer satisfaction from knowledge sharing process (knowledge co-creation) between outside-in and inside-out information.

For the outside-in approach, customer sentiment in each service aspect was retrieved from social media by technology such as text preprocessing, text mining, natural language processing (NLP), and sentiment analysis. In this stage, we developed the sentiment keyword extraction tool named *"Sansarn Tagging tool*" and proposed the novel sentiment analysis tool named "***Multidimensional sentiment cube***" which can provide unlimited viewpoints of analysis by deployed the OLAP concept. This tool can present deeper analysis by across aspects and sub-aspects than the current sentiment analysis approach generally extract in individual aspect level or basic hierarchical platform.

For the inside-out approach, we do the knowledge sharing process using the survey by a questionnaire and interview to collect data in each aspect from specialists, opinion of customers/non-customers in Thai life insurance service. Knowledge co-creation occurs in this step, the expected results are feedback and idea for problem-solving methods from customers, non-customers and expertise in life insurance.

4.1 Sub Research Answer 1

Sub Research Question 1: How to develop analysis tool for extracting customer sentiment knowledge from social media?

This research presents a newish concept of analysis tools named multi-dimensional sentiment cube for extract customer sentiment from social media. It reveals problems from the service process's significant service issues, which are derived from the relationship among life insurance factors under the social CRM concept. It also used unstructured data from user-generated content in social media and then analyzed sentiment to identify problems for support service industries such as life insurance in the social media era. This tools use the traditional concept of OLAP to make use of benefit in sentiment analysis concept. Also this tools can provide very complicated point of views from the across operation features of dimension such as slice-dice, drill-through, drill-down, nested etc. In addition, this system also provides the examples of the process to identify the business process that requires improving due to high negative sentiment from customers using event logs from the cube.

The Example of Customer Sentiment Extraction from Multidimensional Sentiment Cube

In the example of analysis, we presented some analysis scenarios from our multidimensional sentiment cube on the life insurance functions under the concept of social CRM. This tool revealed customer sentiment knowledge from social media as following.

Customer Acquisition:

To provide an analysis example to support CRM's customer acquisition strategy that has the main purpose of approaching new customers and ex-customers. This process requires a positive attitude of the customer and gain high customer satisfaction. In an analysis of the life insurance process, we have to clarify a dissatisfaction point of service.

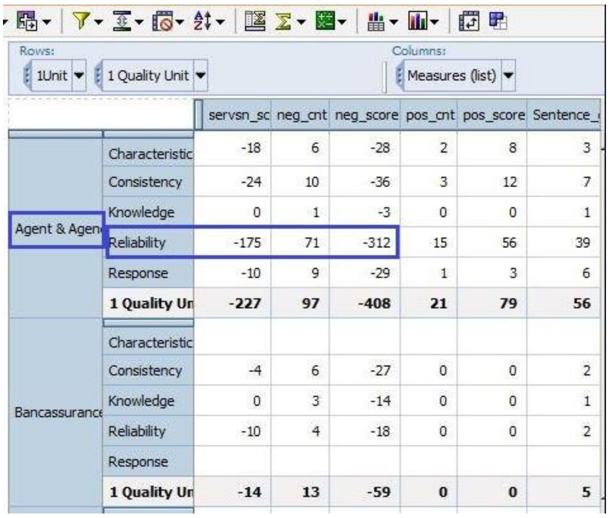
Parent (Top level of aspect) analysis:

{service_manner||service_evaluation} consider to ➔ {service word count & score, negative word count & score, positive word count & score, sentence count}

⇨ Aspect: Service_manner = ‘External service’ has the most negative score of service evaluation which can identify Service Manner group with negative score 14.98% of all negative score.
  - Non-Match, External service, Internal service and Multi-group get 81.89%, 14.98%, 2.64% and 0.5% respectively.

    Child (Sub-aspect) analysis: deep down to find out lower level which related to parent level

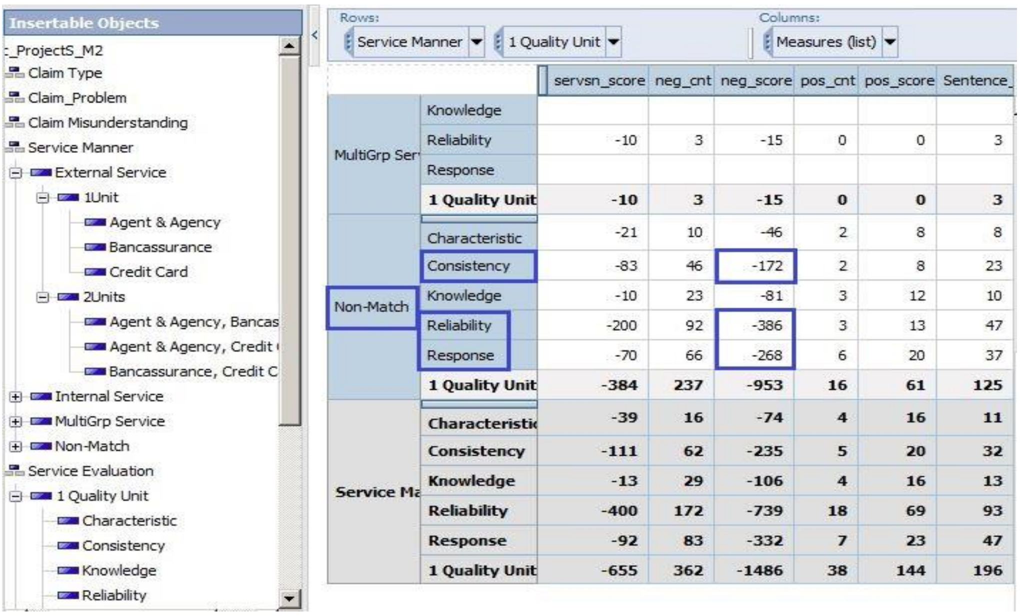
⇨ Sub-aspect: ‘External service’ = ‘Agent & Agency’ has negative score of service evaluation on ‘Reliability’>’Consistency’>’Response’>’Characteristic’>
’Knowledge’ with negative score = 76%, 8.823%, 7.108%, 6.86% and 0.73% respectively. (Figure 26)



		servsn_sc	neg_cnt	neg_score	pos_cnt	pos_score	Sentence_
Agent & Agen	Characteristic	-18	6	-28	2	8	3
	Consistency	-24	10	-36	3	12	7
	Knowledge	0	1	-3	0	0	1
	Reliability	-175	71	-312	15	56	39
	Response	-10	9	-29	1	3	6
	1 Quality Un	**-227**	**97**	**-408**	**21**	**79**	**56**
Bancassurance	Characteristic						
	Consistency	-4	6	-27	0	0	2
	Knowledge	0	3	-14	0	0	1
	Reliability	-10	4	-18	0	0	2
	Response						
	1 Quality Un	**-14**	**13**	**-59**	**0**	**0**	**5**

Figure 26: An analysis example of customer acquisition 1

In customer acquisition, the most dissatisfaction of customer is 'reliability' to agent & agency who is representative to offer new product/plan of life insurance to the customer. There are untrustworthy problems on service touchpoints and issues on inconsistency and irresponsibility, which need to improve to increase customer satisfaction. As we inform about 'Non-match' issue consideration, even if we cannot map clue-terms of 'Service manner' aspect, it can match clue-terms on 'Service evaluation' aspects, and then cube can provide useful analysis view. The results show the service has problems on 'Reliability,' Response,' ' Consistency,' 'Knowledge,' 'Characteristic' respectively with negative score = 25.98%, 18.04%, 11.58%, 5.45%, and 3.1% respectively under Non-match group with a negative score 64.13% from all negative score of Service manner (Figure 27).



		servsn_score	neg_cnt	neg_score	pos_cnt	pos_score	Sentence_
MultiGrp Ser	Knowledge						
	Reliability	-10	3	-15	0	0	3
	Response						
	1 Quality Unit	**-10**	**3**	**-15**	**0**	**0**	**3**
Non-Match	Characteristic	-21	10	-46	2	8	8
	Consistency	-83	46	-172	2	8	23
	Knowledge	-10	23	-81	3	12	10
	Reliability	-200	92	-386	3	13	47
	Response	-70	66	-268	6	20	37
	1 Quality Unit	**-384**	**237**	**-953**	**16**	**61**	**125**
Service Ma	**Characteristic**	**-39**	**16**	**-74**	**4**	**16**	**11**
	Consistency	**-111**	**62**	**-235**	**5**	**20**	**32**
	Knowledge	**-13**	**29**	**-106**	**4**	**16**	**13**
	Reliability	**-400**	**172**	**-739**	**18**	**69**	**93**
	Response	**-92**	**83**	**-332**	**7**	**23**	**47**
	1 Quality Unit	**-655**	**362**	**-1486**	**38**	**144**	**196**

Figure 27: An analysis example of customer acquisition 2

Customer Retention:

We analyzed aspects/sub-aspects related to life insurance operation on the service process to support customer retention strategy. Our example made use of the claim assessment module as below.

Rows: Claim_Problem (... Columns: Measures (list) Conte

	dmprob_cnt	dmprob_score	neg_cnt	neg_scor	pos_cnt	pos_score	Sentence_
Advance Money	26	-66	32	-109	4	16	23
Conceal	92	-283	67	-286	13	48	78
Fax	31	-5	21	-87	8	27	29
Long Service	43	-132	83	-344	7	29	40
Reject Payment	147	-721	216	-887	18	70	142
Advance Money, Conceal	3	-10	1	-3	0	0	1
Advance Money, Fax	5	-18	4	-16	0	0	2
Advance Money, Reject Payment	60	-239	51	-182	9	30	29
Conceal, Fax	2	-4	1	-3	0	0	1
Reject Payment, Conceal	16	-68	18	-76	1	4	8
Reject Payment, Fax	10	-28	9	-34	0	0	5
Advance Money, Reject Payment, Fax	6	-16	3	-9	0	0	2
Advance Money, Reject Payment, Long Servic	3	-11	2	-7	0	0	1
Total	**444**	**-1601**	**508**	**-2043**	**60**	**224**	**361**

Figure 28: An analysis example of customer retention 1

Parent analysis:

{claim_problem} consider to ➔ {claim word count & score, negative word count & score, positive word count & score, sentence count}

⇨ Aspect: Claim problem = 'Reject Payment,' 'Long Service' and 'Conceal' has the high negative score of service evaluation with 43.42%, 16.84%, and 13.99% of total claim problem score respectively. Besides, we found 'Advance Money' has a high associated with 'Reject Payment' in negative sentiment score with 8.91% of the total claim problem score. It means customer complaints about their frequency at the same time with high dissatisfaction.

These results show customers' opinions which are dissatisfied on 'Reject payment.' It means customers expect to get the full money price of a medical bill they already paid to hospital return from an insurance company when they request claim assessment. If they did not meet their expectation, dissatisfaction would occur as a result in Figure 28. As a multidimensional sentiment cube feature, it provides multi-analysis viewpoints to change the other viewpoints in customer retention. For example, Figure 29 shows customers' dissatisfaction with claim assessment in the Rider product group more than the Life product group. Furthermore, there are high negative sentiments on misunderstanding point in claim assessment of customer on the benefit of protection (Misunderstanding in protection). The next problem is the period of claim (Misunderstanding in Period).

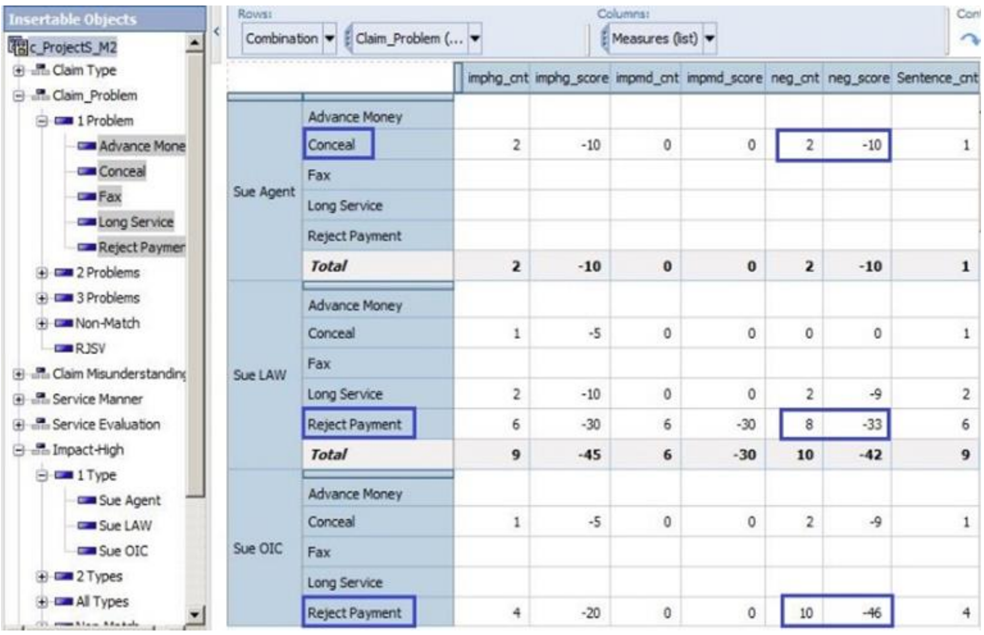
		dmmis_cr	dmmis_score	neg_cnt	neg_scor	pos_cnt	pos_score	Sentence_
LIFE	MisUnderstand in Coverage	7	-29	11	-49	0	0	4
	MisUnderstand in Protection							
	MisUnderstand in Period							
	MisUnderstand in Exception	8	8	12	-53	2	8	8
	Total	**15**	**-21**	**23**	**-102**	**2**	**8**	**12**
Rider	MisUnderstand in Coverage	4	-16	2	-9	0	0	4
	MisUnderstand in Protection	15	-72	18	-70	1	4	15
	MisUnderstand in Period	2	-10	4	-16	1	4	2
	MisUnderstand in Exception	4	4	2	-8	0	0	3
	Total	**25**	**-94**	**26**	**-103**	**2**	**8**	**24**

Figure 29: An analysis example of customer retention 2

Parent analysis:
{product||Insufficient understanding} consider to ➔ {claim misunderstanding count & score, negative word count & score, positive word count & score, sentence count}

⇨ Aspect: product = 'Rider' and sub-aspect = 'Insufficient understanding' has a high negative score has the most negative score on 'Misunderstanding in Protection.'
  - All group of product get negative score on Non-match, Life, Rider, special product and affinity product 61.12%, 19.20%, 18.55%, 1.06% and 0.07% respectively.

Customer Termination:
An analysis example of customer termination, we chose multidimensional cube features to find the customers' sentiment why customers intent to take legal action proceedings against insurance company using two aspects analysis, Impact_High related to Claim_Problem aspect as Figure 30.



		imphg_cnt	imphg_score	impmd_cnt	impmd_score	neg_cnt	neg_score	Sentence_cnt
Sue Agent	Advance Money							
	Conceal	2	-10	0	0	2	-10	1
	Fax							
	Long Service							
	Reject Payment							
	Total	**2**	**-10**	**0**	**0**	**2**	**-10**	**1**
Sue LAW	Advance Money							
	Conceal	1	-5	0	0	0	0	1
	Fax							
	Long Service	2	-10	0	0	2	-9	2
	Reject Payment	6	-30	6	-30	8	-33	6
	Total	**9**	**-45**	**6**	**-30**	**10**	**-42**	**9**
Sue OIC	Advance Money							
	Conceal	1	-5	0	0	2	-9	1
	Fax							
	Long Service							
	Reject Payment	4	-20	0	0	10	-46	4

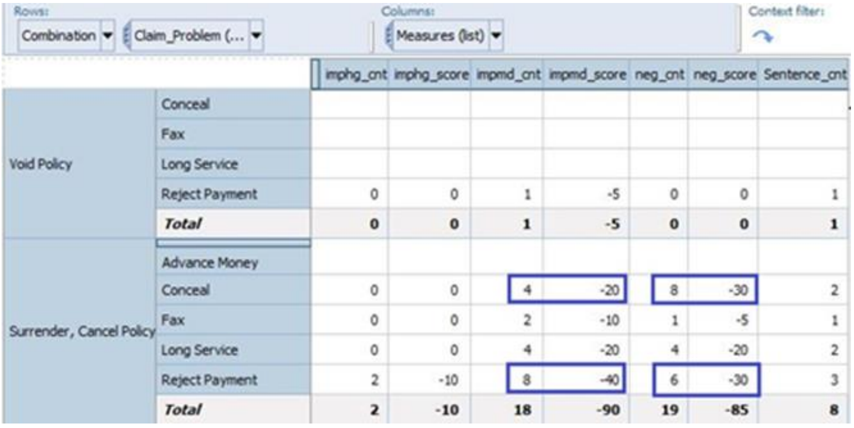
Figure 30: An analysis example of customer termination 1

Child analysis:
{Impact-High||Claim_Problem} consider to➔ {imphg word count & score, negative word count & score, positive word count & score, sentence count}

- ⇨ Sub-Aspect: Impact-High = 'Sue on Law' related to Claim_Problem = 'Reject Payment' has the most negative score.
- ⇨ Sub-Aspect: Impact-High = 'Sue on OIC' related to Claim_Problem = 'Reject Payment'. It means when the insurance company rejects to pay claim assessment return to the customer. This is one of the main reasons which influence to make a petition for the Office of Insurance Commission.
- ⇨ Sub-Aspect: Impact-High = 'Sue on Agent' related to Claim_Problem = 'Conceal' has the most negative score.

It means concealed fraud behavior of service touchpoint influences to accuse agent. An overall negative score of High impact (Exclude non-match) of sue on law, sue on OIC, sue agent, two aspects group are 42.7%,27.8%, 9.83%, and 19.66% respectively.

Also, we use a multidimensional cube on customers' sentiment to find out why customers leave their life insurance contracts. We start to clarify these problems on Impact_Medium from the parent level, which negatively score on surrender & cancel the policy, surrender the policy, invalid policy and cancel the policy (exclude non-match) 87.82%, 7.8%, 3.17%, and 1.22%, respectively. Next, we can analyze the child level, and we found that the 'Surrender' problem with 'Cancel Policy' has a high negative score when the claim problem is 'Reject Payment' and 'Conceal' Figure 31.



Combination	Claim_Problem	imphg_cnt	imphg_score	impmd_cnt	impmd_score	neg_cnt	neg_score	Sentence_cnt
Void Policy	Conceal							
	Fax							
	Long Service							
	Reject Payment	0	0	1	-5	0	0	1
	Total	**0**	**0**	**1**	**-5**	**0**	**0**	**1**
Surrender, Cancel Policy	Advance Money							
	Conceal	0	0	4	-20	8	-30	2
	Fax	0	0	2	-10	1	-5	1
	Long Service	0	0	4	-20	4	-20	2
	Reject Payment	2	-10	8	-40	6	-30	3
	Total	**2**	**-10**	**18**	**-90**	**19**	**-85**	**8**

Figure 31: An analysis example of customer termination 2

- Sub-Aspect: Impact-Medium in ('Surrender', 'Cancel Policy') related to Claim_Problem in ('Reject Payment', ' Conceal') has the most negative score.

4.2 Sub Research Answer 2

Sub Research Question 2: What are the issues relate to customer sentiment in Thai life insurance from knowledge sharing process?

To summarize the results of the customer sentiment extraction related to issues in Thai life insurance regarding customer sentiment, we analyzed customer acquisition, customer retention, and customer termination from multidimensional sentiment cube found four important issues as the following

1. **Service evaluation factors:** The results found that dissatisfaction issues relate to the reliability, knowledge, responsiveness, service consistency, and personality of a service provider.
    - Reliability - ability to provide services agreed to in the contract.
    - Knowledge - the capability to provide efficient knowledge.
    - Responsiveness – readiness and willingness to provide services that respond to customers' needs promptly.
    - Service consistency - enabling customers to feel that the service is reliable and can be trusted.
    - Personality - ability to provide services with courtesy, politeness, and good manners, including effective communications.

2. **Claim settlement**
    - The dissatisfaction factors relevant to claim settlements on social media found from the analysis tool were as follows:
        - Taking a long time to compensate for the policyholder.
        - Receiving lower compensation than the policyholder’s expectation.
        - Policyholder’s claim is rejected (does not receive compensation).
        - The policyholder is required to pay advance health expenses to the hospital.
        - The policyholder does not understand the claim process.
        - Concealing the health record of the policyholder.
        - Policyholder cannot contact service provider for support during the claim process in a timely manner.

3. **Policy cancellation** the results found the policy cancellation related to customer request to cancel the policy on their hands, including policy lapses because a customer did not continue premium payment for many reasons. For example, customers request policy cancellation because of dissatisfaction in service, policy on hand does not meet the fundamental requirement, and the customer's financial status has changed. It cannot continue to pay a premium at the same rate as previous. In addition, the results found that the customers prefer to other life insurance companies rather than current one, and they would like to make a new contract and to reject the current policies. Sometimes, the results found customers gave their opinion about policy cancellation because they require sum assured before the policy maturity period.

4. **Misunderstandings of life insurance principles and regulations** between policyholders and service providers include controversial problems that generate lawsuits or cancellations of life insurance policies, as expressed on social media. The topics found were as follows.
    - Policy regulations related to rejected claims
    - Service provider issues such as agent reliability
    - Lack of knowledge of the service provider

- Lack of prudence by the service provider
- Uncontrolled situations, such as economic instability, changing the total sum assured at policy maturity.

4.3 Sub Research Answer 3

Sub Research Question 3: What kinds of knowledge on customer sentiment have been utilize for knowledge sharing process between service provider and customer?

As ***knowledge*** is what is accumulated from education, research, or experience, including practical ability and skills, understanding, or information gained from experience, gained through hearing, hearing, thinking, or practice in a specific field of study (*https://www.gotoknow.org/posts/396638*). To create customer sentiment knowledge for answering our objectives, we integrate the results of customer sentiment from multidimensional sentiment cube analysis with the results from survey using questionnaire.

Customer sentiment knowledge, which is stated, has two kinds of customer sentiment knowledge sources. It was generated from two ways communication between service providers and customers on social media and the questionnaire.

The first kind of customer sentiment knowledge is 'Outside-in sentiment knowledge.' It includes 'Knowledge of customer sentiment' and 'Knowledge from customer sentiment' as stated in the previous section. Knowledge of customer sentiment is related to customer sentiment, which expresses social media's dissatisfaction with claim settlement policy cancellation and misunderstanding issues. Knowledge from customer sentiment is the dissatisfaction level from negative scoring and reveals the significant of hidden problems with prioritized sentiment scores.

The second kind of customer sentiment knowledge is ‘Inside-out sentiment.’ It is ‘Knowledge for customer sentiment.’ Service provider expertise offers problem-solving methods for dissatisfaction problems, and respondents gave scoring in the questionnaire. The results showed the appreciated problem-solving approach by customers.

1. **Outside-in sentiment knowledge:** *Knowledge of customer sentiment* and *Knowledge from customer sentiment.*

From sub research answer 2, we have got the customer sentiment, which expresses social media's dissatisfaction in service with claim settlement policy cancellation and misunderstanding issues. To integrate and confirm the results between the results from our tools and the survey, we summarize and analyze the answers according to the data's significant factors from the questionnaire using **SPSS**. The questionnaire's evaluation scores used a five-point scale, with one representing least dissatisfied and five representing most dissatisfied. The results of each objective are as follows.

The results from survey in part of Service evaluation factors:

Objective A: to measure customer dissatisfaction with the service evaluation factors of the service providers to prioritize the most important problems.

This study analyzed dissatisfaction based on *lack of reliability-LackReliability*, *lack of knowledge-LackKnowledge, lack of responsiveness–LackResponsiveness, lack of consistency in service-LackConsistency, lack of good personality-LackPersonality* based on life insurance experience, knowledge, and attitudes.

Service Evaluation	Experience		Knowledge			Attitude			Total
	Exp	No_Exp	High	Medium	Less	POS	NEU	NEG	
LackReliability	2.868	2.802	2.870	2.856	2.696	2.804	3.211	2.467	2.832
LackKnowledge	2.874	2.768	2.862	2.796	2.768	2.810	2.868	2.800	2.816
LackResponsiveness	2.898	2.826	2.935	2.814	2.812	2.829	3.237	2.533	2.858
LackConsistency	2.988	2.821	3.014	2.850	2.768	2.872	3.211	2.600	2.896
LackPersonality	2.695	2.729	2.725	2.701	2.725	2.713	2.711	2.733	2.714

Table 13: Measures of customer dissatisfaction on service evaluation factors

Table 13 presents the results of the average mean score values. We found that the most dissatisfied service evaluation factors were "***lack of consistency in service***" and *"**lack of responsiveness**"* based on experience (*2.988, 2.898*), high knowledge (3*.014, 2.935),* and positive and neutral attitudes *(2.872, 2.829, 3.237, and 3.211).* However, the respondents with negative attitudes gave the highest dissatisfied scores to knowledge (2.8) and personality (2.733).

Objective B. to study significant factors that influence service evaluation factors.

This study analyzed dissatisfaction in terms of service evaluation factors consisting of *lack of reliability-LackReliability, lack of knowledge-LackKnowledge, lack of responsiveness–LackResponsiveness, lack of consistency in service-LackConsistency, lack of good personality-LackPersonality* based on demographic characteristics (gender, age, education, occupation, and income) and life insurance experience, knowledge and attitudes. The results are an analysis of the significant factors using **SPSS** and **ANOVA** tests at a significance level *0.05*.

		ANOVA					
Dissatisfaction with service	**Significant factors**		**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
LackReliability	Education	Between Groups	8.945	3	2.982	4.319	0.005
LackReliability	Attitude	Between Groups	8.283	4	2.071	2.983	0.019
LackKnowledge	Education	Between Groups	7.690	4	1.922	2.786	0.026
LackResponsive	Age	Between Groups	17.770	5	3.554	4.250	0.001
LackResponsive	Education	Between Groups	12.969	3	4.323	5.118	0.002

LackResponsive	Attitude	Between Groups	8.894	4	2.223	2.592	0.036
LackConsistency	Age	Between Groups	13.447	5	2.689	2.848	0.015
Avg. of Service dissatisfaction	Age	Between Groups	10.269	5	2.054	4.226	0.001
Avg. of Service dissatisfaction	Education	Between Groups	4.378	3	1.459	2.923	0.034

Table 14: ANOVA results

Lack of reliability:

The results indicate that the lack of reliability does vary significantly according to ***education*** and ***attitude*** as Table 14 shows.

- **Education.** Test of Homogeneity of Variances: Levene = 0.919 and Sig. = 0.432 more than alpha (0.05), so accept H_0, F-test

F-test = 4.319 or Sig. = p-value = 0.005, less than 0.05 (α). The null hypothesis (H_0) is rejected. Therefore, dissatisfaction with reliability does vary significantly according to education.

- **Attitude.** Test of Homogeneity of Variances: Levene = 1.159 and Sig. = 0.329 more than alpha (0.05), so accept H_0, F-test

F-test = 2.983 or Sig. = p-value = 0.019, less than 0.05 (α). The null hypothesis (H_0) is rejected. Therefore, dissatisfaction with reliability does vary significantly according to attitude.

Lack of knowledge:

The results indicate that a lack of knowledge does vary significantly according to ***occupation,*** as Table 14 shows.

- **Occupation.** Test of Homogeneity of Variances: Levene = 0.919 and Sig. = 0.453 more than alpha (0.05,) so accept H_0, F-test

F-test = 2.786 or Sig. = p-value = 0.026, less than 0.05 (α). The null hypothesis (H_0) is rejected. Therefore, dissatisfaction with lack of knowledge does vary significantly according to occupation.

Lack of responsiveness:

The results indicate that a lack of responsiveness does vary significantly according to ***age***, ***education, and attitude,*** as Table 14 shows.

- **Age.** Test of Homogeneity of Variances: Levene = 0.554 and Sig. = 0.735 more than alpha (0.05), so accept H_0, F-test

F-test = 4.25 or Sig. = p-value = 0.001, less than 0.05 (α). The null hypothesis (H_0) is rejected. Therefore, dissatisfaction with responsiveness does vary significantly according to age.

- **Education.** Test of Homogeneity of Variances: Levene = 0.554 and Sig. = 0.753 more than alpha (0.05), so accept H_0, F-test

F-test = 5.118 or Sig. = p-value = 0.002, less than 0.05 (α). The null hypothesis (H_0) is rejected. Therefore, dissatisfaction with responsiveness does vary significantly according to education.

- **Attitude.** Test of Homogeneity of Variances: Levene = 1.061 and Sig. = 0.376 more than alpha (0.05), so accept H_0, F-test

F-test = 2.592 or Sig. = p-value = 0.036, less than 0.05 (α). The null hypothesis (H_0) is rejected. Therefore, the dissatisfaction in responsiveness does vary significantly according to attitude.

Lack of consistency in service:

The results indicate that the lack of service consistency over the long term does vary significantly according to ***age,*** as shown in Table 14.

- **Age.** Test of Homogeneity of Variances: Levene = 0.482 and Sig. = 0.789 more than alpha (0.05), so accept H_0, F-test

F-test = 2.848 or Sig. = p-value = 0.015, less than 0.05 (α). The null hypothesis (H_0) is rejected. Therefore, dissatisfaction with consistency of service does vary significantly according to age.

Average of service dissatisfaction factor:

The results found that the service evaluation factors (average mean score) do vary significantly according to ***age*** and ***education*** as shown in Table 14.

- **Age.** Test of Homogeneity of Variances: Levene = 0.599 and Sig. = 0.701 more than alpha (0.05), so accept H_0, F-test

F-test = 4.226 or Sig. = p-value = 0.001, less than 0.05 (α). The null hypothesis (H_0) is rejected. Therefore, the average mean score of the service evaluation factors does significantly vary according to age.

- **Education.** Test of Homogeneity of Variances: Levene = 1.633 and Sig. = 0.181 more than alpha (0.05), so accept H_0, F-test

F-test = 4.378 or Sig. = p-value = 0.034, less than 0.05 (α). The null hypothesis (H_0) is rejected. Therefore, the average mean score of the service evaluation factors significantly varies according to education.

Causes of policyholder dissatisfaction with claim settlements:

This study analyzed the causes of policyholder dissatisfaction with the claim settlement process, which consisted of *cannot contact service provider–ClmCNContact, get compensation in long term–ClmGetCompLngT, cannot make claim (for that decease)–ClmCNClaim, pay medical expense in advance–Clm_PayMedExpAv, found conceal health record of customer–ClmConceal* based on life insurance experience, knowledge and attitudes.

Service Evaluation	Experience		Knowledge			Attitude			Total
	Exp	No_Exp	High	Medium	Less	POS	NEU	NEG	
ClmCNContact	2.616	2.613	2.710	2.569	2.536	2.589	2.684	3.000	2.615
ClmGetCompLngT	2.778	2.719	2.833	2.665	2.768	2.701	2.868	3.400	2.746
ClmCNClaim	2.539	2.497	2.572	2.491	2.464	2.477	2.684	2.933	2.516

ClmPayMedExpAv	2.988	2.801	3.087	2.749	2.812	2.844	3.026	3.400	2.885
ClmConceal	2.623	2.522	2.667	2.491	2.551	2.539	2.737	2.733	2.567

Table 15: Analysis of causes of policyholder dissatisfaction with claim settlement process

Table 15 presents the results of the average mean score values. We found the same results for three of the analysis bases (life insurance experience, knowledge and attitudes) with the highest policyholder dissatisfaction with claim settlements as first, ***"pay medical expense in advance"*** *(2.988, 2.801, 3.087, 2.749, 2.812, 2.844, 3.026, and 3.4)* and second as ***"get compensation over long term"*** *(2.778, 2.719, 2.833, 2.665, 2.768, 2.701, 2.868, and 3.4)*. This implies that policyholders do not want to reserve money to pay medical expenses in advance or obtain late compensation. Thus, they were dissatisfied with the time taken to receive compensation. However, it is necessary to focus on other dissatisfaction-related problems also. The inability to contact the service provider, concealing health history records, and inability to make a claim received dissatisfaction scores that were similar to the scores of both main problems.

Policyholder dissatisfaction with policy cancellation:

This study analyzed the motive of policyholder dissatisfaction with policy cancellation, consisting of *get unfavorable service–CancUnfovServ, policy does not meet requirement–CancPolNotReg, financial status is changed–CancFinChg, need to change Life insurance company–CancChgCom, need sum assured before policy's maturity period–CancReqSumAs* based on life insurance experience, knowledge, and attitude.

Service Evaluation	Experience		Knowledge			Attitude			Total
	Exp	No_Exp	High	Medium	Less	POS	NEU	NEG	
CancUnfovServ	3.114	3.360	3.210	3.239	3.353	3.237	3.289	3.400	3.249
CancPolNotReq	3.169	3.316	3.210	3.211	3.426	3.232	3.368	3.333	3.250
CancFinChg	3.491	3.495	3.471	3.485	3.559	3.484	3.553	3.533	3.493
CancChgCom	2.729	2.937	2.674	2.946	2.941	2.824	2.947	3.000	2.844
CancRegSumAs	3.156	3.239	3.261	3.145	3.221	3.172	3.237	3.733	3.202

Table 16: Analysis of policyholder dissatisfaction with policy cancellations

Table 16 presents the results of the average mean score values. We found the same results for the three analysis bases (life insurance experience, knowledge, and attitudes, except negative attitude). The highest ranked motive for policy cancellation was ***financial status is changed*** *(3.491, 3.495, 3.471, 3.485, 3.559, 3.484, 3.553 and 3.533)*, followed by ***policy does not meet requirement*** and ***get unfavorable service.***

As most life insurance plans support policyholders over the long-term, that is, more than ten years, twenty years, or whole life, policyholders have to pay premiums over the long term as well. If the policyholder's financial status changes, it becomes difficult to maintain the premium payments, making this a strong motivation to cancel the life insurance policy. Hence, service providers must consider plan adjustments over a long-term service contract. The next motivation referenced for policy cancellation was that policy does not meet the requirement. The agent should consider this problem at the product offer stage before the customer enters the policy contract to ensure the life insurance plan can meet the customer's requirement. The other main motive was to get unfavorable service, which was also shown as a strong motive for policy cancellation.

However, the responses of those who had negative attitudes toward life insurance were quite different from others' reactions. They indicated their most serious concern as "need sum assured before the policy's maturity period." It implies that they might not be satisfied with depositing their money in a life insurance product over the long term and believe such a large sum assured could provide more benefits in other ways.

Reasons for misunderstanding of life insurance principles and regulations:
This study analyzed the **reasons leading to** misunderstandings regarding life insurance principles and regulations between policyholders and service providers. It consists of *misunderstanding regarding many conditions–MisCondition, misunderstanding because of agent informing only about the benefits–MisAgtInfrmBen, misunderstanding owing to agent lack of carefulness–MisAgtLackCare, economic change impacting need for change–MisEcochgBenchg, agent does not have enough knowledge–MisAgtNoKnwldg* based on life insurance experience, knowledge, and attitudes.

Service Evaluation	Experience		Knowledge			Attitude			Total
	Exp	No_Exp	High	Medium	Less	POS	NEU	NEG	
MisCondition	3.635	3.778	3.638	3.665	3.986	3.667	4.000	4.000	3.714
MisAgtInfrmBen	3.612	3.699	3.609	3.581	3.797	3.570	4.026	3.933	3.660
MisAgtLackCare	3.533	3.670	3.507	3.599	3.783	3.579	3.605	4.000	3.609
MisEcochgBenchg	3.263	3.377	3.246	3.317	3.507	3.312	3.342	3.600	3.326
MisAgtNoKnwldg	3.413	3.359	3.384	3.323	3.478	3.352	3.447	3.667	3.383

Table 17: Analysis of misunderstandings regarding life insurance principles and regulations

Table 17 presents the results of the average mean score values. We found the same results for three of the analysis bases (life insurance experience, knowledge, and attitude), namely, that the most common reason for **misunderstandings of life insurance principles and regulations** was the misunderstanding regarding many conditions *(3.635, 3.778, 3.638, 3.665, 3.986, 3.667, 4.0, and 4.0).* The next was a misunderstanding because the agent only informed about the benefits and misunderstanding due to an agent's carefulness. This feedback implies that many life insurance conditions and regulations are easily misunderstood, mainly when agents discuss only the policy's benefits to sell the plan. Moreover, agents may not provide clear information about the conditions of the policy. The final benefits, such as the sum assured based on the provider, dividend, or interest, may vary as economic situations change, leading to further misunderstandings. Hence, life insurance is complicated and requires agents with sufficient knowledge to provide information to customers before the latter sign policy contracts.

2. Inside-out sentiment: *Knowledge for customer sentiment*
Service provider expertise offers problem-solving methods for dissatisfaction problems, and respondents gave scoring in the questionnaire. The results showed the appreciated problem-solving approach by customers.

Objective: to study the appropriate problem-solving methods for dissatisfaction problems.

This study analyzed appropriate problem-solving methods that respondents were optimistic about, consisting of *discussions to solve problems–FixDiscuss, changes in policy types–FixChgPolTyp, consulting with mediating organizations–FixConOrg, issuing premium discounts–FixPremDisc, and offering a gift to policyholders–FixGift* based on life insurance experience, knowledge, and attitudes.

Problem Solving	Experience		Knowledge			Attitude			Total
	Exp	No_Exp	High	Medium	Less	POS	NEU	NEG	
FixDiscuss	3.802	3.652	3.841	3.629	3.696	3.698	3.842	3.867	3.719
FixChgPolTyp	3.569	3.488	3.594	3.467	3.522	3.520	3.684	3.200	3.524
FixConOrg	3.551	3.372	3.580	3.473	3.565	3.533	3.526	3.467	3.529
FixPremDisc	3.108	3.377	3.116	3.287	3.449	3.224	3.368	3.600	3.254
FixGift	3.072	3.242	3.138	3.156	3.246	3.143	3.263	3.400	3.166

Table 18: Analysis of appropriate problem-solving methods for each service problem

All of the base analysis qualities (experience: has/does not have, knowledge: high/medium/low, attitude: positive/neutral/negative) showed the same results with respondents giving the problem-solving method *discussion to solve problems* the highest score: *3.802, 3.652, 3.841, 3.629, 3.696, 3.698, 3.842, and 3.867 out of 5* respectively. The second-highest ranked problem-solving method was *consulting with mediating organizations*. The third-ranked method was allowing *a change in policy type*. However, respondents with negative attitudes held a different perspective, giving discount the premium a score of 3.6 out of 5. Furthermore, although giving gifts to policyholders did not get the top score, it did receive a high score similar to the others.

4.4 Multidimensional Sentiment Cube Evaluation by domain experts and end-users

Reviewing the usage of multidimensional sentiment cube with domain experts who have experience in the Life insurance domain is necessary for understanding their opinion on System Design, System Usage, and System Validation for future improvement. This research interviewed domain experts in Life insurance and the other domain such as telecommunication, retail, and consultant domain. Their work positions consist of management level and general users, and some interviewees have computer skills, and some have less computer skill. The total number of the interviewee is thirteen persons. We evaluate three parts. The first is system design, the second is system usage, and the last is system validation. The results show in the table below.

a. Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?

	Domain Expert		Other Domain		Total
	IT(5)	Non-IT(2)	IT(4)	Non-IT(2)	13
Implement	-	-	-	-	0
Not implement yet	5	2	3	2	12
under study	-	-	1	-	1

Table 19: An interview results 1

We found almost all organizations of interviewees are not start to implement any application related to sentiment analysis, including sentiment cube 92.3%. However, some organization begins to study the possibility of a sentiment system at 7.69%.

b. In your opinion, who is proper to use this system?

	Domain Expert		Other Domain		Total	
	IT(5)	Non-IT(2)	IT(4)	Non-IT(2)	13	%
a. domain expert in insurance	3	2	2	2	9	69.23
b. domain expert in CRM	5	2	2	2	11	84.62
c. business analysis	3	2	3	2	10	76.92
d. IT	2	2	2	2	8	61.54
e. management level	5	2	3	2	12	92.31
f. strategic planner	5	2	2	2	11	84.62
g. user in organization	0	2	0	0	2	15.38
h. general user & customer	0	2	0	1	3	23.08

Table 20: An interview results 2

Interviewees give opinions that this application proper to Management level of company 92.31%. It means the results of this application can help management level make a decision, and the next is a strategic planner and domain expert in CRM on 84.62%. In the interviewees' opinion, they agree to use these results to set up the company's strategies and help support information to the CRM system. They agree that this application is proper to business analysis that provides analysis report to management level on 76.92%. The next is a domain expert in insurance, and IT on 69.23% and 61.54% respectively.

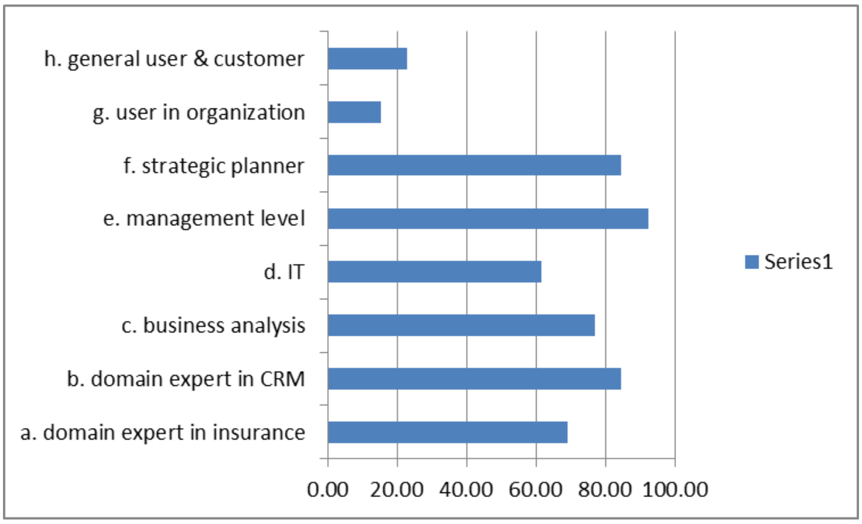


Figure 32: The industry positions who proper to use this system

System Design:

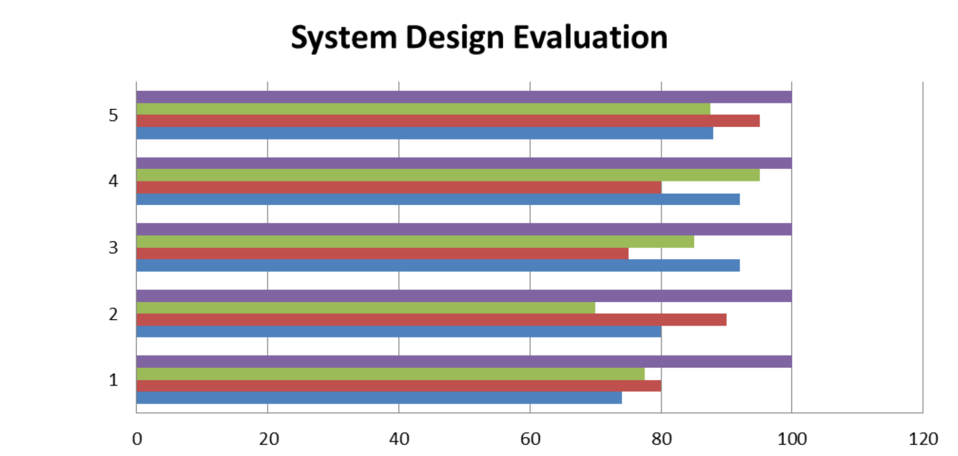
		100%				80%				50%				30%				0%
		D.E.		O.D.		D.E.		O.D.		D.E.		O.D.		D.E.		O.D.		
		IT(5)	NIT(2)	IT(4)	NIT(2)	IT(5)	NIT(2)	IT(4)	NIT(2)	IT(5)	NIT(2)	IT(4)	NIT(2)	IT(5)	NIT(2)	IT(4)	NIT(2)	
1	Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction from social media for CRM?	0	0	1	2	4	2	2	0	1	0	1	0	0	0	0	0	0
2	Do you think CRM concept can help identifying aspect for design?	0	1	1	2	5	1	1	0	0	0	2	0	0	0	0	0	0
3	Do you think process flow can help identifying weak points in current process?	3	1	1	2	2	0	3	0	0	1	0	0	0	0	0	0	0
4	Do you agree with Aspect design in this experiment?	3	0	3	2	2	2	1	0	0	0	0	0	0	0	0	0	0
5	Do you agree with measurement in this experiment by using word cnt and sentiment scoring?	2	3	3	2	3	1	0	0	0	0	1	0	0	0	0	0	0

Table 21: An interview results of system design 1

Remark: D.E. (domain expert) is an interviewee, who is domain expert in life insurance industry, O.D. (other domains) is an interviewee, who works in other domains not related to life insurance, IT is an interviewee, who works relate to information technology, NIT (non-IT) is an interviewee who works not relate to information technology

			Average (%)			
			D.E.		O.D.	
			IT(5)	NIT(2)	IT(4)	NIT(2)
System Design	1	Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction from social media for CRM?	74	80	77.5	100
System Design	2	Do you think CRM concept can help identifying aspect for design?	80	90	70	100
System Design	3	Do you think process flow can help identifying weak points in current process?	92	75	85	100
System Design	4	Do you agree with Aspect design in this experiment?	92	80	95	100
System Design	5	Do you agree with measurement in this experiment by using word cnt and sentiment scoring?	88	95	87.5	100

Table 22: An interview results of system design 2



	1	2	3	4	5
O.D. NIT(2)	100	100	100	100	100
O.D. IT(4)	77.5	70	85	95	87.5
D.E. NIT(2)	80	90	75	80	95
D.E. IT(5)	74	80	92	92	88

Figure 33: The evaluation of system design

The purpose of the interview about system design is to confirm the sentiment analysis, CRM concept, and BPMN – business process flow concept is acceptable for real analysis.

- Sentiment analysis is agreed to support the CRM concept at least 74% by the domain expert (IT) and the most acceptance by other domains (non-IT) 100%.
- CRM concepts are agreed to help identify aspects at least 70% by other domains (IT) and the most accepted by other domains (non-IT) 100%.
- Process flow is agreed for identifying weak points in the current process at least 75% by the domain expert (non-IT) and the most acceptances by other domains (non-IT) 100%.
- Aspect design in this experiment is accepted at least 80% by the domain expert (non-IT) and the most acceptance by other domains (non-IT) 100%.
- The measure of this system use word count and word with sentiment scoring are accepted at least 87.5% by other domain (IT) and the most acceptance by other domains (non-IT) 100%.

Recommendation: There is a recommendation to use emoticon or stickers for sentiment measurement in this system (Interviewee#9). Some interviewees recommend calculating the sentiment scoring variance in statistical terms with sentiment word scoring (Interviewee#5).

System Usage:

		100%				80%				50%				30%				0%
		D.E.		O.D.		D.E.		O.D.		D.E.		O.D.		D.E.		O.D.		
		IT(5)	NIT(2)	IT(4)	NIT(2)	IT(5)	NIT(2)	IT(4)	NIT(2)	IT(5)	NIT(2)	IT(4)	NIT(2)	IT(5)	NIT(2)	IT(4)	NIT(2)	
1	Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction which get from social media?	0	1	3	2	5	1	0	0	0	0	1	0	0	0	0	0	
2	Can the results of data mining rule express the dissatisfaction which get from social media more concretely?	1	0	3	2	3	1	0	0	0	1	1	0	1	0	0	0	

Table 23: An interview results of system usage 1

			Average (%)			
			D.E.		O.D.	
			IT(5)	NIT(2)	IT(4)	NIT(2)
System Usage	1	Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction which get from social media?	80	90	87.5	100
System Usage	2	Can the results of data mining rule express the dissatisfaction which get from social media more concretely?	74	65	87.5	100

Table 24: An interview results of system usage 2

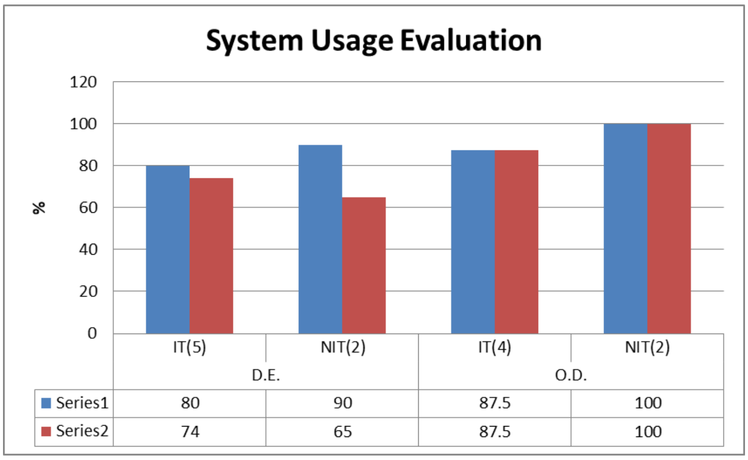


Figure 34: The evaluation of system usage

The purpose to interview about system usage is to make confirmation on the useful of multidimensional sentiment cube and data mining rules.

- Multidimensional sentiment cube can express the relation among features of dissatisfaction which get from social media, interviewees are agreed at least 80% from domain expert (IT) and the most acceptance by other domains (non-IT) 100%.
- Data mining rules can help express dissatisfaction of customer from social media more concretely are agreed at least 65% from domain expert (non-IT) and the most acceptance by other domains (non-IT) 100%.

System Usage	3	Do you prefer to use multidimensional sentiment cube for analysis ?					
			Domain Expert		Other Domain		
			IT(5)	Non-IT(2)	IT(4)	Non-IT(2)	%
		a. Prefer and can analysis by yourself	5	2	4	0	84.62
		b. Prefer but cannot analysis by yourself	0	0	0	2	15.38
		c. Not prefer	0	0	0	0	0.00
System Usage	4	Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve - problems?					
			Domain Expert		Other Domain		
			IT(5)	Non-IT(2)	IT(4)	Non-IT(2)	%
		a. Prefer	4	2	4	2	92.31
		b. Not prefer	1	0	0	0	7.69

Table 25: An interview results of system usage 3

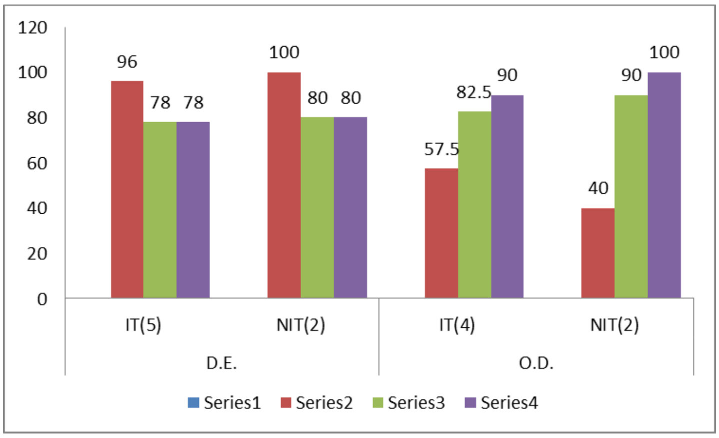


Figure 35: The evaluation on system validation

To survey the difficulty of application usage by asking the preferable usage of cube operation with business knowledge, we found that 84.62% prefer to use this application and confidence to analyze by themselves more than request other people to analyze for them, which have 15.38% (other domain, non-IT).

System usage: What difference between multidimensional sentiment cube usage and data mining rules usage in your opinion?

Answer:
"Multidimensional sentiment cube can answer all of the end-users' questions depending on their requirements; however, data mining rules can change analysis trends and prediction. (*Interviewee#1*)",

"The result of multidimensional sentiment cube and data mining rules is the same; it depends on the requirements of analysis viewpoints; however, the statistical value is significant, mostly negative values. (Interviewee#2)",

"A multidimensional sentiment cube is present in the relational database; however, data mining is for a summary. (*Interviewee#3*)",

"Multidimensional sentiment cube can present in death analysis to serve many analysis requirements. It can represent the abstract value such as satisfaction and dissatisfaction with substantial values; however, data mining is present in terms of substantial values only, yet more concretely. (*Interviewee#4*)",

"Multidimensional sentiment cube is easy for analysis in multi-viewpoints, it shows in term of the relationship of features, so it is relatively easy for understanding and suitable for represent the results of data, data mining rules in particular however it can inform the significant of information (Interviewee#5)".

"Multidimensional sentiment cube shows the relation between features that answer why those problems are occurred, more flexible in analysis term and meet the requirement of analysis ways as end-users want. Data mining rules also provide trend and concrete rules in specific analysis (Interviewee#6)",

"I prefer a multidimensional sentiment cube because it shows figures in all viewpoints which easy to understand and comparable; however, data mining rules are a data conclusion in a feeling of end-users. (Interviewee#7)",

"Multidimensional sentiment cube shows the death analysis viewpoints and data mining rules show the significance of data, so I prefer to use both of information (Interviewee#8)".

"Multidimensional sentiment cube use for analysis in many viewpoints of analysis however data mining rules use in specific analysis (Interviewee#9)",

"Multidimensional sentiment cube use for analysis group or categorize, data mining(association rule) can know the relationship of each aspect with their direction (Interviewee#11)"

System Validation:

		100%				80%				50%				30%				0%			
		D.E.		O.D.		D.E.		O.D.		D.E.		O.D.		D.E.		O.D.		D.E.		O.D.	
		IT(5)	NIT(2)	IT(4)	NIT(2)	IT(5)	NIT(2)	IT(4)	NIT(2)	IT(5)	NIT(2)	IT(4)	NIT(2)	IT(5)	NIT(2)	IT(4)	NIT(2)	IT(5)	NIT(2)	IT(4)	NIT(2)
1	This result is for?																				
	a. Confirm your knowledge	4	2	1	0	1	0	1	1	0	0	1	0	0	0	0	0	0	0	1	1
	b. New knowledge	1	0	3	1	3	2	0	1	1	0	0	0	0	0	1	0	0	0	0	0
2	Can these results help identifying the process which need to improve on dissatisfaction problem?	1	0	2	2	3	2	2	0	1	0	0	0	0	0	0	0	0	0	0	0

Table 26: An interview results of system validation 1

			Average (%)			
			D.E.		O.D.	
			IT(5)	NIT(2)	IT(4)	NIT(2)
System Validation	1	This result is for?				
		a. Confirm your knowledge	96	100	57.5	40
		b. New knowledge	78	80	82.5	90
System Validation	2	Can these results help identifying the process which need to improve on dissatisfaction problem?	78	80	90	100

Table 27: An interview results of system validation 2

To prove the results from application whether they have slightly distorted from life insurance knowledge or not. The interviewee's results reply to this system's results confirms their knowledge at least 40% from other domains (non-IT) interviewees and the most agreement by the domain expert (non-IT) 100%. However, interviewees who are domain experts, both IT and non-IT, are accepted at least 96%.

In addition, all of them can realize new knowledge from the analysis results of cube or data mining rules such as reveal the hidden relationship among features which may be an impact on problems or the number of word count and word sentiment scoring express the level of the relation among features better than previous knowledge from their experience. At least 78% of a domain expert (IT) agrees in new knowledge from this system, and the most agreement from other domains (non-IT) 90%.

Furthermore, these results can help identify the process that needs improvement because of a high negative score from customers. It can analyze using the system's results match on BPMN to determine each problem's related process. At least 80% of a domain expert (non-IT) confirm and the most agreement from other domain interviewees (non-IT) 100%.

			Domain Expert		Other Domain		
System Validation	4	Can Multi-dimensional sentiment cube adapt to another domain?					
			IT(5)	Non-IT(2)	IT(4)	Non-IT(2)	%
		a. Can	5	2	4	2	100
		b. Cannot	0	0	0	0	0

Table 28: An interview results of system validation 3

Next, the results show 100% that this multidimensional sentiment cube concept can adapt to other domains.

Advantage:

Recommend by domain expert (IT) ➔ this system is useful for the organization for analyzing the volume of customers' dissatisfaction, especially in the era of big data generation, which opinion in product or service on social media may express in different results of the traditional way such questionnaire. This system is one tool to make clear pictures of dissatisfaction problems and help reduce complaints to the company site *(Interviewee#1, 4)*. Besides, extracting sentiment from social media is less biased because they can express without informing personal information, so this system is suitable for the industry in the service domain, which does business on the intangible product and requires evaluating satisfaction from the related feature *(Interviewee#3).* And this system also can extract the hidden problem and give feedback before the significant issues occur *(Interviewee#2).*

Recommend by domain expert (non-IT) ➔ this system is useful because creating innovation using technology to develop the analysis method to record in relationship and rules, including business flow, can express a more precise image and easy to understand. Significantly, the measurement of word count and word sentiment score shows a high association to identify positive or negative indicators based on the relevance of data and management theory *(Interviewee#6).* This system can also present the trend and overall pictures of current customers' sentiment level as end-users analysis viewpoints *(Interviewee#7).*

Recommend by other domains (IT) ➔ This system deploys the concept of machine learning to set up rules as a figure for easy explanation to understand in rules of business *(Interviewee#9).* This system also makes use of a social network to be one source to set up a company's strategies *(Interviewee#11).* Also, the measurement in terms of quantitative can bring to analysis in terms of statistical and cube and show detail in descriptive so that this system can support scientific and business management *(Interviewee#8).* Especially, knowledge extraction is also useful and essential *(Interviewee#10).*

Recommend by other domains (non-IT) ➔ this system shows the relationship of key information that is important to provide the reason and impact each other. Cross analysis among features can express the analysis results in in-depth details *(Interviewee#13).* This system can reveal hidden problems in a real business with related features each other *(Interviewee#12).*

Disadvantage:

Recommend by domain expert (IT) ➔ this system is under linguistics limitations because the Thai language is a non-standard language. Language usage is free usage, which may impact word counts and sentiment ranking measurements *(Interviewee#3).* One concern with this system is that the corpus gathered from websites cannot prove whether the story is real *(Interviewee#4).* To provide more statistical value in multidimensional sentiment cube, it is more useful for analysis *(Interviewee#5).* Furthermore, some interviewee has concerned about the high budget to implement this system into a real business *(Interviewee#2).*

Recommend by domain expert (non-IT) ➔ this system requires both skill in cube operation and business knowledge understanding when using this system. Some interviewee recommends revising to store data in a database to protect data lost so this system can work continuously *(Interviewee#6).*

Recommend by other domains (IT) ➔ this system requires high skills, both IT and business knowledge for analysis from this system *(Interviewee#11).* Data tracking such as word segmentation, word sentiment scoring requires double check manually from domain experts and general users after system-generated lexicon or corpus *(Interviewee#9).* Last, an accusation on a competitor in the real world business may occur by posting a fake story on websites that we cannot prove the truth *(Interviewee#10).*

Recommend by other domains (non-IT) ➔ this system concerns data loss from user-generated contents, which occur from language usage by slang words, fashion words, or typos *(Interviewee#13).* To manage the Thai language may take much effort and spend high cost in real implementation *(Interviewee#12).*

Recommendation:

Recommend by domain expert (IT) ➔ the real implementation should increase the step of data source verification by examining the admin of the source of data before crawl data into the system *(Interviewee#4).* For the real usage, if design lexicons to support more deepen down for specific problems, it can reveal more hidden the association among features *(Interviewee#2).* In real usage, a statistical measurement may get better analysis in this system *(Interviewee#5).* Besides, to adapt the other knowledge more than CRM may improve analysis results *(Interviewee#11).*

Recommend by domain expert (non-IT) ➔ It should keep data as a revision to compare the change of analysis results of each period and prepare a backup data plan in real implementation *(Interviewee#6).*

Recommend by other domains (IT) ➔ It will be useful to generate this system with big sample size and collect only believable websites such as has admin to monitor post-blogs. It will be better to prove the truth of comments posted on websites before feeding into the system. It should also set up a recurring task to a re-generated cube with revision data to compare the variance of data. If the variance is low, it can prove the results more concretely *(Interviewee#8).*

Recommend by other domains (non-IT) ➔ in real practice, it requires humans to monitor real sentiment to confirm with the system tracking to reduce mistracking from speaking a language *(Interviewee#13).* So recommend updating the Thai language database (lexicon) more frequently to collect new or modern words useful for semantic learning (*Interviewee#12).*

List of contributors:

1. **Life Insurance domain expert with IT skill**

Interviewee#1

Name:	Capt. Payoon Silagul
Position:	Chief Information Officer (CIO)
Company:	Samart Corporation PCL. (2005 - 2015) & American International Assurance (AIA) (~ 20 years)

Interviewee#2

Name:	Miss Wipa Chareonkitsupat
Position:	Business Unit Manager (Agency) and Financial Advisor
Company:	American International Assurance (AIA)

Interviewee#3

Name:	Miss Woranuch Thanadirake
Position:	Domain Expert in Agent and Agency analysis and CRM
Company:	American International Assurance (AIA)

Interviewee#4

Name:	Mr. Jaturong Soosuk
Position:	Senior solution analysis - Management Information System (IT)
Company:	American International Assurance (AIA)

Interviewee#5

Name:	Miss Rattanavadee Athinantaphan
Position:	Business Analyst consultant
Company:	Thai Cardif Life Insurance, a joint venture company between Thai Life Insurance and Thai Cardif group

2. **Life Insurance domain expert with non-IT**

Interviewee#6

Name:	Miss Suvilai
Position:	
Company:	American International Assurance (AIA)

Interviewee#7

Name:	Miss Vantanee Srisakundacharuk
Position:	Assistant Manager of Customer Service Center
Company:	American International Assurance (AIA)

3. **Other domains with IT skill**

Interviewee#8

Name:	Mr. Anan Derochanawong
Position:	Advisor
Company:	Alliance for Supporting Industries Association (A.S.I.A.)

Interviewee#9

Name:	Mr.Jakkapong Jairuk
Position:	Assistant Manager, Business Support System Di
Company:	Total Access Communication (DTAC)

Interviewee#10

Name:	Mr.Naratip Jamrus

Position:	Hospital Information System Module Add On Software
Company:	Software house

Interviewee#11

Name:	Mr.Suzuki Takeyuki
Position:	Phd.Student
Company:	Japan Advance Science Institute and Technology

4. Other domains with non-IT

Interviewee#12

Name:	Mr.Vichai Janjariyakun
Position:	Vice President of Operation Division
Company:	C.P. All Public company

Interviewee#13

Name:	Mr.Jaturong Kerdrat
Position:	Managing Director
Company:	Getsmart Service

Chapter 5

Conclusions and Recommendations

5.1 Multi-dimensional sentiment cube Comparison with previous works

Many previous works such as M. Castellanos, U. Dayal, C. Gupta, S. Wang, and M. G. Solaco (2012) also introduce the concept of Cubes and hierarchical correlations for analyzing social media data. Multidimensional sentiment cube is closely related to a conceptual modeling framework to define the new requirements set for extracted data from social media using sentiment score associated with these concepts. It mainly presents in terms of a theoretical concept.

However, in this task, the development of the multidimensional sentiment cube is in term of real-world application to study how to adapt the theory of social customer relationship management (social CRM) and Thai life insurance knowledge base to design the semantic hierarchical of association between aspects and sub-aspects including their measurements. Some of the other researchers are interested in a hierarchical structure in sentiment analysis, but they did not experiment on multidimensional cube (item 12, 17). Besides, some papers show multidimensional in unstructured data such as documents data extraction, but they are not related to sentiment analysis (item 11, 14, 15). In term of sentiment analysis, even if many researchers experiment with sentiment analysis widely, however, they (Table 29) presented sentiment analysis concept in term of word-level (item 3, 13), sentence-level (item 3, 13), and aspect-based sentiment analysis (item 3, 8, 9, 10, 12, 16, 18). However, our task presents in terms of a multidimensional cube. Some of them have the primary purpose of increasing sentiment classification accuracy (item 8, 10, 17, 18). Furthermore, many previous works have high intentions in a standard language such as English (item 1-17). This paper utilizes the concept of text mining and natural language processing (NLP) to overcome the obstacle in unique identities (characteristics) of non-standard language such as the Thai language. Additionally, the other previous works in the life insurance domain related to technology are presented in terms of data mining technologies widely; however, it still does not find any papers directly associated with multidimensional sentiment cube.

As a part of knowledge management, the multidimensional sentiment cube provides knowledge analysis to reveal the significant issues that directly impact customers' satisfaction in the Thai life insurance domain. In term of social CRM, this results extraction is a kind of outside-in information for engagement oriented procedure. Then it turns the concept from one-way communication (passive CRM strategy) to be dual-way communication (active CRM strategy) to increase the efficiency of customer acquisition, customer intention, and customer engagement strategy.

		Type			Theory															Domain	
					Sentiment Analysis						Multi dimensional	Text Mining	NLP	Language		DWH	Data Mining	CRM	Knowledge Discovery	Product	Service
		Literature	Application	Theory	Word	Sent.	Aspect-based	Sentiment Classification	Hierarchy	Multidimensional Cube				Eng	Thai						Insurance
1	C. Che-Wei, L. Chin-Tsai and W. Lian-Qing. (2009).		O									O		O		O	O	O			
2	N. E.W.T., X. Li and C. D.C.K. (2009).	O										O		O		O	O	O			
3	H. Minqing and L. Bing. (2004).			O	O	O	O					O	O	O							
4	A.B. Devale and Dr. R.V.Kulkarni. (2012, July).		O											O			O		O		O
5	Shu-Hsien Liao, Ya-Ning Chen and Yu-Yia Tseng. (2009,).													O			O			O	O
6	Yu Yan and Haiying Xie. (2009).	O												O			O		O		O
7	Yongqiang Chen and Leifang Hu. (2005).	O												O			O	O			O
8	Dave, Kushal, Steve Lawrence and David M. Pennock. (2003).		O				O	O				O	O	O						O	
9	Bing Liu, Ding, Xiaowen, and Philip S. Yu. (2008)		O				Lexicon-based						O	O						O	
10	Hu, Minqing and Bing Liu. (2004)		O				O	O				O	O	O						O	
11	R. Frank, T. Olivier, T. Ronan and Z. Gilles. (2007)			O							O			O							
12	K. Suin, Z. Jianwen, O. Alice and L. Shixia. (2013).		O				O		O					O						O	
13	M. Taboada, K. Voll and J. Brooke. (2008		O		O	O							O	O							
14	T. VogelgesangEmail and H.-J. Appelrath. (2015-16)			O							O			O		O	Process		O		
15	Z. Duo, Z. ChengXiang, H. Jiawei (2009)		O								O	O		O		O					
16	Y. Mohd Ridzwan, L. Yuefeng and Z. Jinglan. (2013)		O				O						O	O				O			
17	H. Steven C., Z. Lanqin (2014)		O					O	O					O							Learning
18	C.Haruechaiyasak, A.Kongthon (2010)		O				O	O				O	O		O						Hotel
19	U. Dayal, C. Gupta, M. Castellanos, S. Wang and M. G. Solaco. (2012			O						O				O						O	
20	**A study on Multidimensional Sentiment Cube for social CRM**		O				O			O		O	O		O		O	O	O		O

Table 29: System comparison with previous works

In sum, this research presents a novel method for knowledge extraction by extracting problems from the service issues derived from the relationship among factors under the social CRM concept. It uses unstructured data from user-generated content in social media and then analysis sentiment to identify problems for support service industrial such life insurance in the social media era. And the system also gives an example of the process to identify the business process which requires improving due to get high negative sentiment from a customer using event logs from a cube.

5.2 Conclusion and Recommendations of Multidimensional Sentiment Cube

The dimensions and measurement design is important for the analysis efficiency of multidimensional sentiment cube. The design is deeply details of aspects and sub-aspects to present in the hierarchical of cube dimensions related to business process function gains more benefits in the real analysis. Multidimensional sentiment cube can express the hidden issues from the association among multi-aspects, multi-sub-aspects on their measurements. And cube properties also can handle the missing values tracking on a hierarchical structure in a more upper level of child and parent level.

Nevertheless, this system's drawback is time-consuming in lexicon preparation to present high accuracy in the semantic meaning on the process of text mining and NLP. Even if this system utilizes Sansarn tagging tools to develop the specific domain lexicon, it still requires more knowledge and technology to improve lexicon construction with high accuracy on semantic expression automatically. It helps to reduce the developing time of this system. Another useful recommendation from the interview process that gets from domain experts about the trustworthiness of data sources is that we utilize data from social media such as critiqued weblogging. Crawling the main corpus from trustable websites is highly recommended to gain more reliable results that extract from a multidimensional sentiment cube.

5.3 Conclusion and Recommendations of an Analytic framework of Customer Sentiment

We present the results from an analytical framework, which shows how social media can be used as a beneficial means for problem-solving, using the concept of social CRM. As such, we extracted critical words from the web blog pages to explore customer dissatisfaction with service factors expressed through negative feedback.

This study of Thai life insurance identified service dissatisfaction issues extracted from social media via the sentiment extraction tool, which were then evaluated by responses to our questionnaire. The questionnaire results provided more details, as it helped rank the problems and their possible causes according to the classifications of the different respondent qualifications. Moreover, the sentiment extraction tool can be useful for automatic tracking. As a real-time system, it could be helpful in detecting problems promptly, as they are posted. However, the system’s effectiveness depends on its ability to learn Thai and interpret it accurately.

It is crucial to improve the ethical knowledge of agents for increasing customer satisfaction in Thai life insurance service and their tracking and follow up of long-term service contracts. Such efforts will provide consistent attention to policyholders and will help increase customer satisfaction. The claim system should consider adjustments to reduce prepaid medical expenses and return compensation faster.

As most policy cancellations are due to a policyholder's change in financial status, suggestions include establishing a long-term policy support plan, which helps in premium payment adjustment or the possibility of changing plans. Most misunderstandings arise from the policyholder and the agents' complex conditions, which both the policyholder and the agents find difficult to understand; suggestions include focusing on better education for agents and customers and providing more information with the contract and more excellent honesty regarding the conditions.

The respondents with negative life insurance attitudes often held opinions that differed from others, such as needing the money before its maturity date, as their most important policy cancellation motive. The implication may be that they prefer to spend their money rather than receive the sum assured at the end of the policy. They also gave a high score to agents not taking good care of customers as a major reason leading to misunderstandings.

One limitation of this study is the use of Google documents for our questionnaire distribution, thereby limiting the respondents to internet users and mobile communications. In addition, many respondents resided in large cities in Thailand. For more varied opinions, it is necessary to expand the respondent's sample to various careers and areas, including those without access to the internet. In future studies, we will enhance our sentiment extraction tool's capacity by increasing the underlying knowledge base through lexicon expansion and attempt to collect information from more diverse respondents.

References

1. A.B. Devale and Dr. R.V.Kulkarni. (2012, July). Application of data mining techniques in life insurance, International Journal of data mining&knowledge management process Vol.2, No.4.
2. A.Bolt and W.M. van der Aalst. (2015). Multidimensional Process Mining Using Process Cubes. in International Conference on Enterprise, Business-Process and Information Systems Modeling (CAISE 2015).
3. Abhijit Raorane and R.V.Kulkarni. (2011, August), Data mining techniques: A source for consumer behavior analysis, International Journal of Database Management Systems,Vol.3,No.3.
4. A.Kawtrakul, M.Suktarachan, P.Varasai, A State of the Art of Thai Language Resources and Thai Language Behavior Analysis and Modeling, COLING '02 Proceedings of the 3rd workshop on Asian language resources and international standardization - Volume 12.
5. A.Kawtrakul, T.Sukvakree and J.Charoensuk. (2005). Eelementary discourse unit segmentation for Thai using discource clues and syntactic information. SNLP, Bangkok.
6. Bakar, O., Soykan, Y. & Acar, B., (2018). Measuring Students' Knowledge Level of Life Insurance: A case on the Students of the Department of Insurance and Risk Management at the Dumlupinar University, Turkey. Journal of Educational and Social Research, 8(1), pp. 75-82.
7. Banea, Carmen, Rada Mihalcea, Janyce Wiebe, and Samer Hassan. (2008). Multilingual subjectivity analysis using machine translation. Proc. of the 2008 empirical methods in natural language processing, 127–135.
8. Basaula, D., (2017). Customers Satisfaction towards Life Insurance Claim Settlement in Nepal. Journal of Interdisciplinary Studies, Volume 6.
9. B.Heerschop, F.Goossen and A.Hogenboom. (2011). Polarity Analysis of Texts using Discourse Structure, CIKM’11, Glasgow, Scotland, UK, October 24-28.
10. B. Liu. (2012 May). Sentiment Analysis and Opinion Mining, Morgan & Claypool.
11. B.Liu, Ding, Xiaowen, and Philip S. Yu. (2008). A holistic lexicon-based approach to opinion, WSDM’08 Proc. of the International Conference on web search and mining, 231-240.
12. B. Navid and S. Muhammad. (2012). Customer relationship management constructs under social networks towards customers’ retention," Australian Journal of Basic and Applied Sciences, vol. 6(7), no. ISSN 1991-8178, pp. 271-282.
13. Business Process Modeling Notation (BPMN). (November 2010). TechTarget [Online]. Available: http://searchcio.techtarget.com/definition/Business-Process-Modeling-Notation.
14. C.Apte, E.Grossman, E.Pednault, B.Rosen, F.Tipu, and B.White. (1998, March 31). Insurance risk modeling using data mining technology. IBM T.J. Watson, RC 21314 (94531).
15. C. Che-Wei, L. Chin-Tsai and W. Lian-Qing. (2009). Mining the text information to optimizing the customer relationship management. Expert systems with applications, vol. 36, pp. 1433-1443.

16. C.Haruechaiyasak, A. Kongthon, P. Palingoon and C. Sangkeettrakarn. (2010). Constructing Thai Opinion Mining Resource: A Case Study on Hotel Reviews. in Proceedings of the 8th Workshop on Asian Language Resources, Beijing, China.
17. C.Haruechaiyasak, A.Kongthon, P. Palingoon and K.Trakultaweekoon.(2013 October 14). S-Sense: A sentiment analysis framework for social media sensing. IJCNLP 2013 Workshop on Natural Language Processing for Social Media (SocialNLP), pages 6–13, Nagoya, Japan.
18. C. Haruechaiyasak and N. E. a. C. T. C. (NECTEC). (2010). NLP, Information Retrieval and Text Mining. [Online].
19. Chidanand Apte, Edna Grossman, Edwin P.D.Pednault, Barry K.Rosen, Fatch A. Tipu, and Brian White. (1999). Probabilistic estimation-based data mining for discovering insurance risk, IBM T.J. Watson research center, IBM Research Report RC-21483, IEEE 1094-7167/99.
20. Chompuphan, P., (2014). Factors influencing the purchase of a life insurance policy of AIA Company Limited in Bangkok, Bangkok, Thailand: Dhurakit Pundit University.
21. Coviello, A. & Trapani, G. D., (2012). The Customer Satisfaction in the Insurance Industry. SSRN https://ssrn.com/abstract=2144684.
22. Dave, Kushal, Steve Lawrence and David M. Pennock. (2003). Mining the peanut gallery: opinion extraction and semantic classification of product reviews. Proceedings of the 12th international conference on World Wide Web, 519–528.
23. Davenport, T. H. & Prusak, L., (1998). Working Knowledge: How Organizations Manage What They Know. In: Working Knowledge : How Organizations Manage What They Know. Boston, Massachusetts: Harvard Business School Press.
24. Dennis Wegener and Stefan Ruping. (2010). On Integrating Data Mining into Business Processes., In Proceedings of PADD99.
25. Ding, Xiaowen, Bing Liu and Philip S. Yu. (2008). A holistic lexicon-based approach to opinion, Proc. of the int. conf. on web search and mining, 231–240.
26. Dr Usha Thakur, ATS research and Vinay Kumar Patri, BFSI solutions, HCL technology of Chennai. (2010 April). CRM Trend in insurance, HCL Technologies Ltd.
27. Eui-Hong (Sam) Han, George Karypis and Vipin Kumar. (2012). Min-Apriori:An algorithm for finding association rules in data with continues attributes, Expert systems with application39,11303- 11311.
28. E.W.T Ngai, Li Xiu, D.C.K. Chau. (2009). Application of data mining techniques in customer relationship management: A literature review and classification, Expert systems application 36, 2592 – 2602.
29. Hidaka, K. (2006). Trends in services sciences in Japan and abroad. Quarterly Review 19 pp. 35-47.
30. Hu Guan, Jingyu Zhou and Minyi Guo. (2009). A Class-Feature-Centroid Classifier for Text Categorization, International World Wide Web Conference Committee (IW3C2), WWW 2009, April 20– 24.

31. H. Minqing and L. Bing. (2004). Mining Opinion Features in Customer Reviews in American Association for Artificial Intelligence (www.aaai.org).
32. Hu, Minqing and Bing Liu. (2004). Mining summarizing customer reviews. Proceedings of the SIGKDD international conference on Knowledge discovery and data mining, 168–177.
33. H. Steven C., Z. Lanqin, K. Vive and Kinshuk. (2014). Multi-Dimensional Sentiment Classification in Online Learning Environment," in IEEE Sixth International Conference on Technology for Education.
34. http://siam-processmining.blogspot.com/p/process-mining.html.
35. I. K. Center. InfoSphere Data Architect 9.1.0 Hierarchies. IBM, [Online]. Available:https://www.ibm.com/support/knowledgecenter/en/SS9UM9_9.1.0/com.ibm.datatools.dimensional.ui.doc/topics/c_dm_hierarchy.html.
36. Inoue, T., September (2014). Life Insurance Marketing Consumer Behavior Approach, Tokyo, Japan: Oriental Life Insurance Cultural Development Center.
37. Jianbing Xiahou and Yang Mu. (2010, April). Analysis on application of data mining upon rate making and risk analysis in insurance. Computer Engineering and Technology (ICCET), 2010 2nd International Conference on (Volume:2).
38. Jim Spohrer and Paul P. Maglio. (2008, May – June). The emergence of Service Science: Toward systematic service innovations to accelerate co-creation of value. IBM Almaden research center, Production and Operations Management Society, Vol. 17, No. 3, 238–246.
39. J. S. Ribeiro and A. M. Weijters. (2011). Event Cube: Another Perspective on Business Processes. in OTM 2011: On the Move to Meaningful Internet Systems: OTM 2011.
40. Kannan, M., (2018). Customer Satisfaction towards life insurance corporation with special reference to Chennai City. International Journal of Development and Sustainability, 7(4), pp. 1264-1272.
41. K. Ravi and V. Ravi. (2015 November). A survey on opinion mining and sentiment analysis. Knowledge-Based System, vol. 89, no. 15, pp. 14-46.
42. K. Suin, Z. Jianwen, O. Alice and L. Shixia. (2013). "A Hierarchical Aspect-Sentiment Model for Online Reviews," in Association for the Advancement of Artificial Intelligence (www.aaai.org).
43. Kuhlemeyer, G. A. & Allen, G. H., January (1999). Consumer Satisfaction With Life Insurance: A Benchmarking Survey. Journal of Financial Counseling and Planning, Volume 10, pp. 35-48.
44. K. Yuusuk. (2010, Jun 10). Principle of Claim Assessment. [Online]. Available: https://www.gotoknow.org/posts/365518.
45. L. Fangtao, H. Minlie and Z. Xiaoyan. (2010). Sentiment Analysis with Global Topics and Local Dependency. in Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10).
46. Malthouse, E. C. et al., (2013). Managing Customer Relationship in the Social Media Era: Introduction the Social CRM House. Journal of Interactive Marketing, pp. 270-280.

47. M. C. Edward, H. Michael and S. Bernd. (2013). Managing Customer Relationships in the Social Media Era: Introducing the Social CRM house, vol. 27, Journal of Interactive Marketing, pp. 270-280.
48. Minqing Hu and Bing Lui. (2011, April). Feature-Based Opinion Mining and Summarization.ppt
49. Micah, A Basic Guide to Choosing a Life Insurance Plan. (2012,February 20). From www.money- choice.org/basic-life-insurance-plan-guide.
50. M. Taboada, K. Voll and J. Brooke. (2008, December 2). Extracting sentiment as a function of discourse structure and topicality. Technical Report 2008-20 School of Computing Science, Simon Fraser University, Canada.
51. M.Wiegand, A.Balahur, Andr´es Montoyo, B.Roth and D.Klakow. (2010). A Survey on the Role of Negation in Sentiment Analysis, Saarland University Saarbr¨ucken, Germany and University of Alicante Alicante, Spain.
52. M. Svetlana, N. Thomas and S. H. Marc. (2007 November 5-9), Multidimensional Data Modeling for Business Process Analysis. International Conference on Conceptual Modeling, vol. 26.
53. Muhammad, M. B., Muliadi & Sarminah, S., (2015). The Moderating effect of Perception and Attitude on Purchase Intention towards foreign MNES: Evidence from Indonesian Banking and Retail Industry. International Journal of Business and Management Studies, p. 279–298.
54. M. v. Rosing, S. White, F. Cummins and H. d. Man. (2005). The Complete Business Handbook (Body of Knowledge from Process Modeling to BPM) Volume1. in Business Process Model and Notation—BPMN, LEADing Practice ApS. Published by Elsevier Inc., pp. 429-452.
55. Nan-Chen Hsieh and Kuo-Chung Chu. (2009). Enhancing customer behavior analysis by data mining techniques, International Journal of information and management science, 39-53.
56. NECTEC, LEXiTRON: an online Thai-English Electronic Dicitonary, Retrieved http://lexitron.nectec.or.th, (1995).
57. NECTEC, "sansarn.com," NECTEC. (2014). [Online]. Available: http://sansarn.com/corpus-tagging/health/tag.php.
58. N.E.W.T., X. Li and C. D.C.K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. Expert systems with applications, vol. 36, pp. 2592-2602.
59. Nguyen, H. T., Nguyen, H., Nguyen, N. D. & Phan, A. C., (2018). Determinants of Customer Satisfaction and Loyalty in Vietnamese Life-Insurance Setting. Sustainability, Volume 10, p. 1151.
60. N.Sucharittham, T.Theeramunkong, C.Haruechaiyasak, B.T.Ho, and H.C.Dam. (2013). Data Mining for Life Insurance Knowledge Extraction: A Survey. In Asian Conference on Information Systems 2013, 111–118.
61. N.Sucharittham, C.Haruechaiyasak and T.Theeramunkong. and H.C.Dam (2014). Exploiting Linguistic Knowledge for Sentiment Classification in Online Opinions related to Life Insurance Topic in Thailand," in The Asian Conference on Information Systems, Vietnam. 239 – 246.
62. N.Sucharittham, T.H.Nguyen, T.Theeramunkong, and H.C.Dam (2020). Customer’s Sentiment in Thai Life Insurance. International Journal of Management, Business, and Economics 7: 55–74.

63. N.Tongtep and T.Theeramunkong. (2013). Multi-Stage Automatic NE and PoS Annotation using Pattern-Based and Statistical-Based Techniquies for Thai Corpus Construction. in Asian Language Resources coolocated wtih the 5th International Joint Conference on Natural Language Processing (IJCNLP).
64. Nitzan and I. a. B. Libai. (2011). Social Effects on Customer Retention. Journal of Marketing, vol. 75, pp. 24-38.
65. P. C. Verhoef, R. Venkatesan, L. McAllister, E. C. Malthouse, M. Kraft and S. Ganesan. (2010). CRM in Data Rich Multi-channel Retailing Environments: A Review and Future Research Directions. Journal of Interactive Marketing, Vols. 24,2, pp. 124-37.
66. Phromsuwan, V., (2011). Consumers' Attitude and Total Satisfaction on Buying Life Insurance by Telemarketing at Siam Commercial New York Life Insurance Company in Bangkok Metropolitan Area, Bangkok, Thailand: Master of Business Administration Degree in Marketing at Srinakharinwirot University.
67. Pooja Dave. (2009, February 26). Let Life Insurance Riders Drive your Coverage. From www.investopia.com/articles/pf/07/life_insurance_rider.asp.
68. Rainardi, V., (2015). Data Warehousing and Business Intelligence. [Online] Available at: https://dwbi1.wordpress.com/2015/02/09/populating-fact-tables/
69. R. Frank, T. Olivier, T. Ronan and Z. Gilles. (2007, November 5-9). A Conceptual Model for Multidimensional Analysis of Documents. in Conceptual Modeling - ER 2007, 26th International Conference on Conceptual Modeling, Auckland, New Zealand.
70. Ruta, U., Daiva, Z. & Virginija, C., (2008). The Model of Creation of Customer Relationship Management (CRM) System. The Economic Conditions of Enterprise Functioning of Engineering Economics, No.3(ISSN 1392-2785), p. 58.
71. Samarasinghe, U. et al., (2018). Impact of Service Quality on Customer Satisfaction: With Reference To Life Insurance Services in Sri Lanka. Journal of Business and Management (IOSR-JBM), pp. 55-62.
72. Shu-Hsien Liao, Ya-Ning Chen and Yu-Yia Tseng. (2009, July). Mining demand chain knowledge of life insurance market for new product development. Expert Systems with Applications: An International Journal archive Volume 36 Issue 5, 9422-9437.
73. Siddiqui, M. H. & Sharma, T. G., (2010). Analyzing customer satisfaction with service quality in life insurance services. Journal of Targeting, Measurement and Analysis for Marketing, Volume 18, pp. 221-238.
74. Singchoo, P., (2012). GotoKnow: Customer Relationship Management (CRM). [Online] Available at: https://www.gotoknow.org/posts/496000.
75. Sogunro, A. & Abiola, B., (2014). Measuring Customer Satisfaction on Life Insurance Product: Case Study: Lagos State, Nigeria. International Journal. Manag. Bus. Res., 4(2), pp. 73-80.
76. Spohrer, J., Maglio, P., Bailey, J. and Gruhl, D. (2007, January). Steps toward a science of service systems. IEEE Computer Magazine 71-77.
77. S. Sudprasert and A. Kawtrakul. (2005). Thai Word Segmentation based on Global and Local Unsupervised Learning, Bangkok: Human Language Technology Laboratory (HLT), Computer Engineering Human Language

Technology, National Electronics and Computer Technology Center and Department of Computer Engineering.

78. Subashini, S. & Velmurugan, D., (2016). Policyholders Satisfaction of Life Insurance Products with Reference to Coimbatore District. International Journal of Business and Management Invention, 5(4), pp. 28-33.

79. TechTarget, n.d. Managing the customer lifecycle: customer retention and development. [Online] Available at: http://media.ttgtmedia.com/searchCRM/

80. Trusov, Michael, E. Randolph, Bucklin and K. Pauwels. (2009). Effects of Word-of-Mouth Versus Traditional Marketing: Finding from an Internet Social Networking Site. Journal of Marketing, Vols. 73,5, pp. 90-102.

81. T. Vogelgesang and H.-J. Appelrath. (2015). Multidimensional Process Mining with PMCube Explorer. BPM.

82. T. VogelgesangEmail and H.-J. Appelrath. (2016 July 22). PMCube: A Data-Warehouse-Based Approach for Multidimensional Process Mining. in International Conference on Business Process Management (BPM 2015.

83. T. Wilson, J. Wiebe and P. Hoffmann. (2005, October 6-8). Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis," in Association for Computational Linguistics, Vancouver, British Columbia, Canada.

84. U. Dayal, C. Gupta, M. Castellanos, S. Wang and M. G. Solaco. (2012, October 15-18). Of Cubes, DAGs and Hierarchical Correlations: A Novel Conceptual Model for Analyzing Social Media Data. in Volume 7532 of the series Lecture Notes in Computer Science, 31st International Conference ER 2012, Florence, Italy.

85. U. Dayal, "Sentiment Cube". Patent US20120197950 A1, 2 August 2012.

86. U. Ruta, Z. Daiva and C. Virginija. (2008). The Model of Creation of Customer Relationship Management (CRM) System. The Economic Conditions of Enterprise Functioning of Engineering Economics, vol. No.3, no. ISSN 1392-2785, p. 58.

87. V. Doorn, Jenny, K. Lemon, V. Mittal, S. Nass, D. Pick, P. Pirner and P. C. Verhoef. (2010). Customer Engagement Behavior: Theoretical Foundations and Research Directions," Journal of Service Research, Vols. 13,3, pp. 253-66.

88. V. Kumar, J. Andrew Petersen and R. P. Leone. (2010). Driving Profitability by Encouraging Customer Referrals: Who, When and How. Journal of Marketing, vol. 74, pp. 1-17.

89. V. Rainardi. (2015, February 9). Data Warehousing and Business Intelligence. Blog at WordPress.com.[Online].Available: *https://dwbi1.wordpress.com/2015/02/09/populating-fact-tables/.*

90. V. Thomas, K. Georg, R.-M. Stefanie Rinderle- and A. H.-Jurgen. (2016). Multidimensional Process Mining: Questions, Requirements, and Limitations. in Proceedings of the CAiSE'16 Forum at the 28th International Conference on Advanced Information System Engineering, Ljubljana, Slovenia, 13-17.6.

91. Weinberg, B. D. Berger and P. D. Berger. (2011). Connected Customer Lifetime Value Impact Social Media. Journal of Direct, Data and Digital Marketing Practice, Vols. 12, 4, pp. 328-44.

93. Wiebe, Janyce and Ellen Riloff. (2005). Creating subjective and objective sentence classifiers from unannotated texts, Proc. of Conference on Intelligent Text Processing and Computational Linguistics, 486–497.

94. Wongwiratchit, W., (2015). Customer Satisfaction in Life Insurance in Bangkok, Thailand, Bangkok, Thailand: Faculty of Commerce and Accountancy, Thammasat University.
95. X. Liu, K. Tang, J. Hancock and J. Ha. (2013 April 2-5). A Text Cube Approach to Human, Social and Cultural Behavior in the Twitter Stream. in the 6th International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction (SBP 2013), Washington, DC, USA.
96. Y. Mohd Ridzwan, L. Yuefeng and Z. Jinglan. (2013). Integration of Sentiment Analysis into Customer Relational Model: The Importance of Feature Ontology and Synonym," in The 4th International Conference on Electrical Engineering and Informatics (ICEEI 2013), Malaysia.
97. Yongqiang Chen and Leifang Hu. (2005). Study on Data Mining Application in CRM System Based on Insurance Trade, Proceedings of the 7th international conference on E-commerce.
98. Yu Yan and Haiying Xie. (2009). Research on the application of data mining technology in insurance informatization. Ninth International Conference on Hybrid Intelligent Systems.
99. Z. Duo, Z. ChengXiang, H. Jiawei, S. Ashok and O. Nikunj. (2009 December). Topic Modeling for OLAP on Multidimensional Text Databases: Topic Cube and its Applications. Journal of Statistical Analysis and Data Mining - Best of SDM'09, vol. 2, no. 5-6, pp. 378-395.

Appendices

Appendix A

Multidimensional Sentiment Cube Evaluation Questionnaire List

Interviewee#1 name:		Capt. Payoon Silagul			(Domain Expert with IT skill)		
Position:		Chief Information Officer (CIO)					
Company:		Samart Corporation PCL. (2005 - 2015) & American International Assurance (AIA) (~ 20 years)					
Contact:		payoongsak@gmail.com					
1	Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?						
	a.	not implement					
✓	b.	under study					
✓	c.	never heart before					
2	In your opinion, who is proper to use this system?						
✓	a.	domain expert		✓	a1.	Insurance	
				✓	a2.	CRM	
				✓	a3.	Business Analysis	
				✓	a4.	IT	
✓	b.	Management					
✓	c.	Strategic planner					
	d.	User in organization					
	e.	General User					
System Design:							
3	Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction from social media for CRM?						
✓	Yes	100%	✓80%	50%	30%	0%	
4	Do you think CRM concept can help identifying aspect for design?						
✓	Yes	100%	✓80%	50%	30%	0%	
5	Do you think process flow can help identifying weak points in current process?						
✓	Yes	100%	✓80%	50%	30%	0%	
6	Do you agree with Aspect design in this experiment?						
✓	Yes	100%	✓80%	50%	30%	0%	
7	Do you agree with measurement in this experiment by using word cnt and sentiment scoring?						
✓	Yes	100%	✓80%	50%	30%	0%	
8	Do you have any recommendations?						
	Do not have						
System Usage:							
9	Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction - which get from social media?						
✓	Yes	100%	✓80%	50%	30%	0%	
10	Can the results of data mining rule express the dissatisfaction which get from social media more concretely?						
✓	Yes	100%	80%	50%	✓30%	0%	
11	Which measurement do you prefer to use for analysis ?						
	a.	word cnt	b.	score cnt	✓ c.	both	
12	Do you prefer to use multidimensional sentiment cube for analysis ?						
✓	a.	Prefer and can analysis by yourself					
	b.	Prefer but cannot analysis by yourself					
	c.	Not prefer					

13	Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?										
	a.	Prefer									
✓	b.	Not prefer									
14	What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?										
	Multidimensional sentiment cube can answer in all of questions of end-users depend on their requirements -										
	however data mining rules can analysis trend and prediction.										
System Validation:											
15	This result is for?										
✓	a.	Confirm your knowledge			100%	✓80%	50%	30%	0%		
✓	b.	New knowledge			100%	✓80%	50%	30%	0%		
16	Can these results help identifying the process which need to improve on dissatisfaction problem?										
✓	Yes	100%	✓80%	50%	30%	0%					
	Can Multi-dimensional sentiment cube adapt to another domain?										
17	a.	Can									
	b.	Cannot									
18	Advantage										
	This system is useful for organization for analysis the volume of customers' dissatisfaction -										
	especially in the era of big data generation which opinion in product or service on social media -										
	may express in different results of traditional way such questionnaire. This system is one tool to -										
	make clear pictures of dissatisfaction problem and can help reduce complaint to company.										
19	Disdvantage										
	-										
20	Recommendation										
	-										

Interviewee#2 name		Miss Wipa Chareonkitsupat			(Domain Expert with IT skill)						
Position:		Business Unit Manager (Agency) and Financial Advisor									
Company:		American International Assurance (AIA)									
Contact:		(+6685)110-5905									
1	Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?										
✓	a.	not implement									
	b.	under study									
✓	c.	never heart before									
2	In your opinion, who is proper to use this system?										
✓	a.	domain expert			a1.	Insurance					
				✓	a2.	CRM					
					a3.	Business Analysis					
					a4.	IT					
✓	b.	Management									
✓	c.	Strategic planner									
	d.	User in organization									
	e.	General User									
System Design:											
3	Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction from -										
	social media for CRM?										
✓	Yes	100%	✓80%	50%	30%	0%					
4	Do you think CRM concept can help identifying aspect for design?										
✓	Yes	100%	✓80%	50%	30%	0%					

5 Do you think process flow can help identifying weak points in current process?

✓ Yes | ✓ 100% | 80% | 50% | 30% | 0%

6 Do you agree with Aspect design in this experiment?

✓ Yes | ✓ 100% | 80% | 50% | 30% | 0%

7 Do you agree with measurement in this experiment by using word cnt and sentiment scoring?

✓ Yes | ✓ 100% | 80% | 50% | 30% | 0%

8 Do you have any recommendations?

Do not have

System Usage:

9 Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction - which get from social media?

✓ Yes | 100% | ✓ 80% | 50% | 30% | 0%

10 Can the results of data mining rule express the dissatisfaction which get from social media more concretely?

✓ Yes | 100% | ✓ 80% | 50% | 30% | 0%

11 Which measurement do you prefer to use for analysis ?

a. word cnt | b. score cnt | ✓ c. both

12 Do you prefer to use multidimensional sentiment cube for analysis ?

✓ a. Prefer and can analysis by yourself
b. Prefer but cannot analysis by yourself
c. Not prefer

13 Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?

✓ a. Prefer
b. Not prefer

14 What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?

The result of multidimensional sentiment cube and data mining rules is the same, it depends on the requirements - of analysis viewpoints however the statistical value is important especially negative values.

System Validation:

15 This result is for?

✓ a. Confirm your knowledge | ✓100% | 80% | 50% | 30% | 0%
b. New knowledge | 100% | 80% | ✓50% | 30% | 0%

16 Can these results help identifying the process which need to improve on dissatisfaction problem?

✓ Yes | 100% | 80% | ✓50% | 30% | 0%

17 Can Multi-dimensional sentiment cube adapt to another domain?

a. Can
b. Cannot

18 Advantage

This system can extract the hidden problem and give feedback before the big problems occurs.

19 Disdvantage

To implement this system may require high budget into real business.

20 Recommendation

For the real usage, if design lexicons to support more deepen down for specific problems, it can reveal more hidden the association among features.

Interviewee#3 name:		Miss Woranuch Thanadirake				(Domain Expert with IT skill)					
Position:		Domain Expert in Agent and Agency analysis and CRM									
Company:		American International Assurance (AIA)									
Contact:		Woranucht2@gmail.com									
1	Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?										
✓	a.	not implement									
	b.	under study									
	c.	never heart before									
2	In your opinion, who is proper to use this system?										
✓	a.	domain expert			a1.	Insurance					
				✓	a2.	CRM					
					a3.	Business Analysis					
					a4.	IT					
✓	b.	Management									
✓	c.	Strategic planner									
	d.	User in organization									
	e.	General User									
System Design:											
3	Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction -										
	from social media for CRM?										
✓	Yes	100%	✓80%	50%	30%	0%					
4	Do you think CRM concept can help identifying aspect for design?										
✓	Yes	100%	✓80%	50%	30%	0%					
5	Do you think process flow can help identifying weak points in current process?										
✓	Yes	✓100%	80%	50%	30%	0%					
6	Do you agree with Aspect design in this experiment?										
✓	Yes	✓100%	80%	50%	30%	0%					
7	Do you agree with measurement in this experiment by using word cnt and sentiment scoring?										
✓	Yes	✓100%	80%	50%	30%	0%					
8	Do you have any recommendations?										
	Do not have										
System Usage:											
9	Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction -										
	which get from social media?										
✓	Yes	100%	✓80%	50%	30%	0%					
10	Can the results of data mining rule express the dissatisfaction which get from social media more concretely?										
✓	Yes	✓100%	80%	50%	30%	0%					
11	Which measurement do you prefer to use for analysis ?										
	a.	word cnt	b.	score cnt	✓	c.	both				
12	Do you prefer to use multidimensional sentiment cube for analysis ?										
✓	a.	Prefer and can analysis by yourself									
	b.	Prefer but cannot analysis by yourself									
	c.	Not prefer									

13 Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?

✓ a. Prefer

b. Not prefer

14 What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?

Multidimensional sentiment cube present in term of relational database however data mining is for summary.

System Validation:

15 This result is for?

✓	a.	Confirm your knowledge	✓100%	80%	50%	30%	0%
✓	b.	New knowledge	✓100%	80%	50%	30%	0%

16 Can these results help identifying the process which need to improve on dissatisfaction problem?

✓	Yes	✓100%	80%	50%	30%	0%

17 Can Multi-dimensional sentiment cube adapt to another domain?

✓ a. Can

b. Cannot

18 Advantage

To extract sentiment from social media is less bias because they can express without inform personal information - so this system is suitable for industry in service domain which make business on intangible product and require - to evaluate satisfaction from related feature.

19 Disdvantage

This system is under linguistics limitations because Thai language is non-standard language and the usage of language - is freely usage which may impact on word counts and sentiment ranking measurements.

20 Recommendation

-

Interviewee#4 name	Mr Jaturong Soosuk	(Domain Expert with IT skill)
Position:	Senior solution analysis - Management Information System (IT)	
Company:	American International Assurance (AIA)	
Contact:	(+662) 638-6375	

1 Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?

✓ a. not implement

b. under study

c. never heart before

2 In your opinion, who is proper to use this system?

✓	a.	domain expert	✓	a1.	Insurance
			✓	a2.	CRM
			✓	a3.	Business Analysis
				a4.	IT
✓	b.	Management			
✓	c.	Strategic planner			
	d.	User in organization			
	e.	General User			

System Design:

3 Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction - from social media for CRM?

✓	Yes	100%	✓80%	50%	30%	0%

4 Do you think CRM concept can help identifying aspect for design?

✓	Yes	100%	✓80%	50%	30%	0%

5		Do you think process flow can help identifying weak points in current process?						
	✓	Yes	100%	✓ 80%	50%	30%	0%	
6		Do you agree with Aspect design in this experiment?						
	✓	Yes	100%	✓ 80%	50%	30%	0%	
7		Do you agree with measurement in this experiment by using word cnt and sentiment scoring?						
	✓	Yes	100%	80%	✓ 50%	30%	0%	
8		Do you have any recommendations?						

System Usage:

9		Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction - which get from social media?						
	✓	Yes	100%	✓ 80%	50%	30%	0%	
10		Can the results of data mining rule express the dissatisfaction which get from social media more concretely?						
	✓	Yes	100%	80%	✓ 50%	30%	0%	
11		Which measurement do you prefer to use for analysis ?						
		a.	word cnt	b.	score cnt	✓ c.	both	
12		Do you prefer to use multidimensional sentiment cube for analysis ?						
	✓	a.	Prefer and can analysis by yourself					
		b.	Prefer but cannot analysis by yourself					
		c.	Not prefer					
13		Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?						
	✓	a.	Prefer					
		b.	Not prefer					
14		What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?						
		Multidimensional sentiment cube can present in death analysis to serve many analysis requirements and can represent the abstract value such satisfaction, dissatisfaction with substantial values in the same time, However, data mining present in term of substantial values only however more concretely.						

System Validation:

15		This result is for?							
	✓	a.	Confirm your knowledge	100%	✓ 80%	50%	30%	0%	
	✓	b.	New knowledge	100%	80%	✓ 50%	30%	0%	
16		Can these results help identifying the process which need to improve on dissatisfaction problem?							
	✓	Yes	100%	✓80%	50%	30%	0%		
17		Can Multi-dimensional sentiment cube adapt to another domain?							
	✓	a.	Can						
		b.	Cannot						
18		Advantage							
		This system is useful for organization for analysis the volume of customers' dissatisfaction.							
19		Disdvantage							
		The corpus which was gathered from websites cannot prove whether the story is real.							
20		Recommendation							
		In the real implementation should increase the step of data source verification by examine in admin of source - of data before crawl data into system.							

Interviewee#5 name: Miss Rattanavadee Athinantaphan (Domain Expert with IT skill)
Position: Business Analyst consultant
Company: Thai Cardif Life Insurance, a joint venture company between Thai Life Insurance and Thai Cardif group
Contact: (+6686) 327-9757

1 Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?

✓	a.	not implement			
	b.	under study			
✓	c.	never heart before			

2 In your opinion, who is proper to use this system?

✓	a.	domain expert	✓	a1.	Insurance
			✓	a2.	CRM
			✓	a3.	Business Analysis
			✓	a4.	IT
✓	b.	Management			
✓	c.	Strategic planner			
	d.	User in organization			
	e.	General User			

System Design:

3 Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction - from social media for CRM?

✓	Yes	100%	80%	✓50%	30%	0%

4 Do you think CRM concept can help identifying aspect for design?

✓	Yes	100%	✓80%	50%	30%	0%

5 Do you think process flow can help identifying weak points in current process?

✓	Yes	✓100%	80%	50%	30%	0%

6 Do you agree with Aspect design in this experiment?

✓	Yes	✓100%	80%	50%	30%	0%

7 Do you agree with measurement in this experiment by using word cnt and sentiment scoring?

✓	Yes	✓100%	80%	50%	30%	0%

8 Do you have any recommendations?
Recommend to calculate about the sentiment scoring variance in statistical term with sentiment word scoring.

System Usage:

9 Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction - which get from social media?

✓	Yes	100%	✓80%	50%	30%	0%

10 Can the results of data mining rule express the dissatisfaction which get from social media more concretely?

✓	Yes	100%	✓80%	50%	30%	0%

11 Which measurement do you prefer to use for analysis ?

a.	word cnt	b.	score cnt	✓	c.	both

12 Do you prefer to use multidimensional sentiment cube for analysis ?

✓	a.	Prefer and can analysis by yourself
	b.	Prefer but cannot analysis by yourself
	c.	Not prefer

13		Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?						
	✓	a.	Prefer					
		b.	Not prefer					
14		What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?						
		Multidimensional sentiment cube is easy for analysis in multi-viewpoints, it shows in term of relationship of features -						
		so it is quite easy for understanding and suitable for represent the results of data, data mining rules is very specific -						
		however it can inform the significant of information.						
System Validation:								
15		This result is for?						
	✓	a.	Confirm your knowledge	✓100%	80%	50%	30%	0%
	✓	b.	New knowledge	100%	✓80%	50%	30%	0%
16		Can these results help identifying the process which need to improve on dissatisfaction problem?						
	✓	Yes	100%	✓80%	50%	30%	0%	
17		Can Multi-dimensional sentiment cube adapt to another domain?						
	✓	a.	Can					
		b.	Cannot					
18		Advantage						
		This system show the trend of customers' sentiment in multi-feature association.						
19		Disdvantage						
		Providing more statistical value in multidimensional sentiment cube make more useful for analysis.						
20		Recommendation						
		In the real usage, to provide the statistical measurement may get better analysis in this system.						

Interviewee#6 name			Miss Suvilai Pangkavibul		(Domain Expert with non-IT)			
Position:			Customer Service Center Manager					
Company:			American International Assurance (AIA)					
Contact:			Suvilai Pangkavibul@aia.co.th (+882) 638-7556					
1		Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?						
	✓	a.	not implement					
		b.	under study					
		c.	never heart before					
2		In your opinion, who is proper to use this system?						
	✓	a.	domain expert	✓	a1.	Insurance		
				✓	a2.	CRM		
				✓	a3.	Business Analysis		
				✓	a4.	IT		
	✓	b.	Management					
	✓	c.	Strategic planner					
	✓	d.	User in organization					
	✓	e.	General User					
System Design:								
3		Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction -						
		from social media for CRM?						
	✓	Yes	100%	✓80%	50%	30%	0%	
4		Do you think CRM concept can help identifying aspect for design?						
	✓	Yes	✓100%	80%	50%	30%	0%	

5 Do you think process flow can help identifying weak points in current process?

✓ Yes ✓100% 80% 50% 30% 0%

6 Do you agree with Aspect design in this experiment?

✓ Yes 100% ✓80% 50% 30% 0%

7 Do you agree with measurement in this experiment by using word cnt and sentiment scoring?

✓ Yes ✓100% 80% 50% 30% 0%

8 Do you have any recommendations?

Do not have

System Usage:

9 Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction - which get from social media?

✓ Yes ✓100% 80% 50% 30% 0%

10 Can the results of data mining rule express the dissatisfaction which get from social media more concretely?

✓ Yes ✓100% 80% 50% 30% 0%

11 Which measurement do you prefer to use for analysis ?

a. word cnt b. score cnt ✓ c. both

12 Do you prefer to use multidimensional sentiment cube for analysis ?

✓ a. Prefer and can analysis by yourself

b. Prefer but cannot analysis by yourself

c. Not prefer

13 Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?

✓ a. Prefer

b. Not prefer

14 What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?

Multidimensional sentiment cube show relation between features which answer the reason why that problems - are occurred, more flexible in analysis term and meet requirement of analysis ways as end-user want. Data mining rules also provide trend and concrete rules in specific analysis.

System Validation:

15 This result is for?

✓ a. Confirm your knowledge ✓100% 80% 50% 30% 0%

✓ b. New knowledge 100% ✓80% 50% 30% 0%

16 Can these results help identifying the process which need to improve on dissatisfaction problem?

✓ Yes ✓100% 80% 50% 30% 0%

17 Can Multi-dimensional sentiment cube adapt to another domain?

✓ a. Can

b. Cannot

18 Advantage

This system is useful because creating new innovation using technology for developing the analysis method - to record in relationship and rules including business flow which can express more clearly image and easily - to understand especially, the measurement of word count and word sentiment score show highly association - to identify positive or negative indicator based on the relevant of data and management theory.

19 Disdvantage

This system requires both skill in cube operation and business knowledge understanding when use this system.

20 Recommendation

Recommends make revision to store data in database in order to protect data lost so this system can work continuously It should keep data as revision to compare change of analysis results of each period and should prepare back up - data plan in real implementation.

Interviewee#7 name:		Miss Vantanee Srisakundacharuk			(Domain Expert with non-IT)		
Position:		Assistant Manager of Customer Service Center					
Company:		American International Assurance (AIA)					
Contact:		(+6685) 022-2512					
1	Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?						
✓	a.	not implement					
	b.	under study					
	c.	never heart before					
2	In your opinion, who is proper to use this system?						
✓	a.	domain expert		✓	a1.	Insurance	
				✓	a2.	CRM	
				✓	a3.	Business Analysis	
				✓	a4.	IT	
✓	b.	Management					
✓	c.	Strategic planner					
✓	d.	User in organization					
✓	e.	General User					
System Design:							
3	Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction -						
	from social media for CRM?						
✓	Yes	100%	✓80%	50%	30%	0%	
4	Do you think CRM concept can help identifying aspect for design?						
✓	Yes	100%	✓80%	50%	30%	0%	
5	Do you think process flow can help identifying weak points in current process?						
✓	Yes	100%	80%	✓50%	30%	0%	
6	Do you agree with Aspect design in this experiment?						
✓	Yes	100%	✓80%	50%	30%	0%	
7	Do you agree with measurement in this experiment by using word cnt and sentiment scoring?						
✓	Yes	100%	✓80%	50%	30%	0%	
8	Do you have any recommendations?						
	Do not have						
System Usage:							
9	Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction -						
	which get from social media?						
✓	Yes	100%	✓80%	50%	30%	0%	
10	Can the results of data mining rule express the dissatisfaction which get from social media more concretely?						
✓	Yes	100%	80%	✓50%	30%	0%	
11	Which measurement do you prefer to use for analysis ?						
	a.	word cnt	b.	score cnt	✓	c.	both
12	Do you prefer to use multidimensional sentiment cube for analysis ?						
✓	a.	Prefer and can analysis by yourself					
	b.	Prefer but cannot analysis by yourself					
	c.	Not prefer					

13	Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?									
✓	a.	Prefer								
	b.	Not prefer								
14	What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?									
	I prefer multidimensional sentiment cube because it show figures in all viewpoints which easy to understand and -									
	comparable however data mining rules is a data conclusion in feeling of end-users.									
System Validation:										
15	This result is for?									
✓	a.	Confirm your knowledge		✓100%	80%	50%	30%	0%		
✓	b.	New knowledge		100%	✓80%	50%	30%	0%		
16	Can these results help identifying the process which need to improve on dissatisfaction problem?									
✓	Yes	100%	✓80%	50%	30%	0%				
17	Can Multi-dimensional sentiment cube adapt to another domain?									
✓	a.	Can								
	b.	Cannot								
18	Advantage									
	This system can present the trend and overall pictures of currently of customers' sentiment level as end-users -									
	analysis viewpoints.									
19	Disdvantage									
	-									
20	Recommendation									
	-									

Interviewee#8 name		Mr. Anan Derochanawong			(Other Domain with IT skill)					
Position:		Advisor								
Company:		Alliance for Supporting Industries Association (A.S.I.A.)								
Contact:		(+662) 712-1729								
1	Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?									
✓	a.	not implement								
	b.	under study								
	c.	never heart before								
2	In your opinion, who is proper to use this system?									
✓	a.	domain expert		✓	a1.	Insurance				
				✓	a2.	CRM				
					a3.	Business Analysis				
					a4.	IT				
	b.	Management								
	c.	Strategic planner								
	d.	User in organization								
	e.	General User								
System Design:										
3	Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction -									
	from social media for CRM?									
✓	Yes	100%	80%	✓50%	30%	0%				
4	Do you think CRM concept can help identifying aspect for design?									
✓	Yes	100%	80%	✓50%	30%	0%				

5	Do you think process flow can help identifying weak points in current process?								
✓	Yes	100%	✓ 80%	50%	30%	0%			
6	Do you agree with Aspect design in this experiment?								
✓	Yes	✓ 100%	80%	50%	30%	0%			
7	Do you agree with measurement in this experiment by using word cnt and sentiment scoring?								
✓	Yes	✓ 100%	80%	50%	30%	0%			
8	Do you have any recommendations?								
	Do not have								
System Usage:									
9	Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction -								
	which get from social media?								
✓	Yes	✓ 100%	80%	50%	30%	0%			
10	Can the results of data mining rule express the dissatisfaction which get from social media more concretely?								
✓	Yes	✓ 100%	80%	50%	30%	0%			
11	Which measurement do you prefer to use for analysis ?								
	a.	word cnt	b.	score cnt	✓	c.	both		
12	Do you prefer to use multidimensional sentiment cube for analysis ?								
✓	a.	Prefer and can analysis by yourself							
	b.	Prefer but cannot analysis by yourself							
	c.	Not prefer							
13	Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?								
✓	a.	Prefer							
	b.	Not prefer							
14	What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?								
	Multidimensional sentiment cube show death analysis viewpoints and data mining rules show the significant of data -								
	so I prefer to use both of information.								
System Validation:									
15	This result is for?								
✓	a.	Confirm your knowledge			✓ 100%	80%	50%	30%	0%
✓	b.	New knowledge			✓ 100%	80%	50%	30%	0%
16	Can these results help identifying the process which need to improve on dissatisfaction problem?								
✓	Yes	✓ 100%	80%	50%	30%	0%			
17	Can Multi-dimensional sentiment cube adapt to another domain?								
✓	a.	Can							
	b.	Cannot							
18	Advantage								
	The measurement in term of quantitative can bring to analysis in term of statistical and cube also show detail -								
	in descriptive so this system can support both of scientific and business management.								
19	Disdvantage								
	-								
20	Recommendation								
	It will be good if generate this system with big sample size and collect only the believable websites -								
	such has admin to monitor post-blogs and it will be better to prove the truth of comments which are posted -								
	in websites before feed into system. In addition, it should set up recurring task to re-generated cube with revision data -								
	in order to compare the variance of data. If the variance is less, it can prove the results more concretely.								

Interviewee#9 name:		Mr.Jakkapong Jairak				(Other Domains with IT skill)	
Position:		Assistant Manager, Business Support System Division					
Company:		Total Access Communication (DTAC)					
Contact:		jakkapong.jairak@dtac.co.th					
1	Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?						
✓	a.	not implement					
	b.	under study					
	c.	never heart before					
2	In your opinion, who is proper to use this system?						
✓	a.	domain expert			a1.	Insurance	
					a2.	CRM	
				✓	a3.	Business Analysis	
				✓	a4.	IT	
✓	b.	Management					
✓	c.	Strategic planner					
	d.	User in organization					
	e.	General User					
System Design:							
3	Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction -						
	from social media for CRM?						
✓	Yes	100%	✓80%	50%	30%	0%	
4	Do you think CRM concept can help identifying aspect for design?						
✓	Yes	100%	✓80%	50%	30%	0%	
5	Do you think process flow can help identifying weak points in current process?						
✓	Yes	100%	✓80%	50%	30%	0%	
6	Do you agree with Aspect design in this experiment?						
✓	Yes	100%	✓80%	50%	30%	0%	
7	Do you agree with measurement in this experiment by using word cnt and sentiment scoring?						
✓	Yes	100%	80%	✓50%	30%	0%	
8	Do you have any recommendations?						
	Recommend to use emoticon or stickers for sentiment measurement include into this system						
System Usage:							
9	Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction -						
	which get from social media?						
✓	Yes	100%	80%	✓50%	30%	0%	
10	Can the results of data mining rule express the dissatisfaction which get from social media more concretely?						
✓	Yes	100%	80%	✓50%	30%	0%	
11	Which measurement do you prefer to use for analysis ?						
	a.	word cnt	b.	score cnt	✓	c.	both
12	Do you prefer to use multidimensional sentiment cube for analysis ?						
✓	a.	Prefer and can analysis by yourself					
	b.	Prefer but cannot analysis by yourself					
	c.	Not prefer					

13		Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?									
	✓	a.	Prefer								
		b.	Not prefer								
14		What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?									
		Multidimensional sentiment cube use for analysis in many viewpoints of analysis however data mining rules use -									
		in specific analysis.									
System Validation:											
15		This result is for?									
	✓	a.	Confirm your knowledge		100%	✓80%	50%	30%	0%		
	✓	b.	New knowledge		100%	80%	50%	✓30%	0%		
16		Can these results help identifying the process which need to improve on dissatisfaction problem?									
	✓	Yes	100%	✓80%	50%	30%	0%				
17		Can Multi-dimensional sentiment cube adapt to another domain?									
	✓	a.	Can								
		b.	Cannot								
18		Advantage									
		This system deploy the concept of machine learning to set up rules as a figure for easily explanation to -									
		understand in rules of business.									
19		Disdvantage									
		Data tracking such as word segmentation, word sentiment scoring requires double check manually -									
		from domain experts and general users after system generated lexicon or corpus.									
20		Recommendation									
		-									

Interviewee#10 name		Mr.Naratip Jamrus			(Other Domains with IT skill)		
Position:		Hospital Information System Module Add On Software					
Company:		Software house					
Contact:		(+6693) 183-9617					
1		Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?					
✓	a.	not implement					
	b.	under study					
	c.	never heart before					
2		In your opinion, who is proper to use this system?					
✓	a.	domain expert			a1.	Insurance	
					a2.	CRM	
				✓	a3.	Business Analysis	
					a4.	IT	
✓	b.	Management					
✓	c.	Strategic planner					
	d.	User in organization					
	e.	General User					
System Design:							
3		Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction -					
		from social media for CRM?					
✓	Yes	✓100%	80%	50%	30%	0%	
4		Do you think CRM concept can help identifying aspect for design?					
✓	Yes	✓100%	80%	50%	30%	0%	

5	Do you think process flow can help identifying weak points in current process?								
✓	Yes	✓100%	80%	50%	30%	0%			
6	Do you agree with Aspect design in this experiment?								
✓	Yes	✓100%	80%	50%	30%	0%			
7	Do you agree with measurement in this experiment by using word cnt and sentiment scoring?								
✓	Yes	✓100%	80%	50%	30%	0%			
8	Do you have any recommendations?								
	Do not have								
System Usage:									
9	Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction - which get from social media?								
✓	Yes	✓100%	80%	50%	30%	0%			
10	Can the results of data mining rule express the dissatisfaction which get from social media more concretely?								
✓	Yes	100%	✓80%	50%	30%	0%			
11	Which measurement do you prefer to use for analysis ?								
	a.	word cnt	b.	score cnt	✓	c.	both		
12	Do you prefer to use multidimensional sentiment cube for analysis ?								
✓	a.	Prefer and can analysis by yourself							
	b.	Prefer but cannot analysis by yourself							
	c.	Not prefer							
13	Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?								
✓	a.	Prefer							
	b.	Not prefer							
14	What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?								
	-								
System Validation:									
15	This result is for?								
✓	a.	Confirm your knowledge			100%	✓80%	50%	30%	0%
✓	b.	New knowledge			✓100%	80%	50%	30%	0%
16	Can these results help identifying the process which need to improve on dissatisfaction problem?								
✓	Yes	✓100%	80%	50%	30%	0%			
17	Can Multi-dimensional sentiment cube adapt to another domain?								
✓	a.	Can							
	b.	Cannot							
18	Advantage								
	Knowledge extraction which get from system is useful and important.								
19	Disdvantage								
	An accusation on competitor in the real world business may occur by post the fake story into websites - which we cannot prove the truth.								
20	Recommendation								
	-								

Interviewee#11 name:		Mr.Suzuki Takeyuki				(Other Domain with IT skill)					
Position:		Phd.Student									
Company:		Japan Advance Science Institute and Technology									
Contact:		t_suzuki@jaist.ac.jp									
1	Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?										
	a.	not implement									
✓	b.	under study									
	c.	never heart before									
2	In your opinion, who is proper to use this system?										
✓	a.	domain expert		✓	a1.	Insurance					
				✓	a2.	CRM					
				✓	a3.	Business Analysis					
					a4.	IT					
✓	b.	Management									
✓	c.	Strategic planner									
	d.	User in organization									
	e.	General User									
System Design:											
3	Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction -										
	from social media for CRM?										
✓	Yes	100%	✓80%	50%	30%	0%					
4	Do you think CRM concept can help identifying aspect for design?										
✓	Yes	100%	80%	✓50%	30%	0%					
5	Do you think process flow can help identifying weak points in current process?										
✓	Yes	100%	✓80%	50%	30%	0%					
6	Do you agree with Aspect design in this experiment?										
✓	Yes	✓100%	80%	50%	30%	0%					
7	Do you agree with measurement in this experiment by using word cnt and sentiment scoring?										
✓	Yes	✓100%	80%	50%	30%	0%					
8	Do you have any recommendations?										
	Do not have										
System Usage:											
9	Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction -										
	which get from social media?										
✓	Yes	✓100%	80%	50%	30%	0%					
10	Can the results of data mining rule express the dissatisfaction which get from social media more concretely?										
✓	Yes	✓100%	80%	50%	30%	0%					
11	Which measurement do you prefer to use for analysis ?										
	a.	word cnt	b.	score cnt	✓	c.	both				
12	Do you prefer to use multidimensional sentiment cube for analysis ?										
✓	a.	Prefer and can analysis by yourself									
	b.	Prefer but cannot analysis by yourself									
	c.	Not prefer									

13	Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?								
✓	a.	Prefer							
	b.	Not prefer							
14	What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?								
	Multidimensional sentiment cube use for analysis group or categorize, data mining(association rule) can know -								
	relationship of each aspect with their direction.								
System Validation:									
15	This result is for?								
	a.	Confirm your knowledge		100%	80%	50%	30%	✓0%	
✓	b.	New knowledge		✓100%	80%	50%	30%	0%	
16	Can these results help identifying the process which need to improve on dissatisfaction problem?								
✓	Yes	100%	✓80%	50%	30%	0%			
17	Can Multi-dimensional sentiment cube adapt to another domain?								
✓	a.	Can							
	b.	Cannot							
18	Advantage								
	This system also makes use of social network to be one source to set up company's strategies.								
19	Disdvantage								
	This system requires highly skill both of IT and business knowledge for analysis from this system.								
20	Recommendation								
	-								

Interviewee#12 name		Mr.Vichai Janjariyakun			(Other Domains with non-IT)				
Position:		Vice President of Operation Division							
Company:		C.P. All Public company							
Contact:		(+6691) 789-6999							
1	Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?								
✓	a.	not implement							
	b.	under study							
	c.	never heart before							
2	In your opinion, who is proper to use this system?								
✓	a.	domain expert		✓	a1.	Insurance			
				✓	a2.	CRM			
				✓	a3.	Business Analysis			
				✓	a4.	IT			
✓	b.	Management							
✓	c.	Strategic planner							
	d.	User in organization							
	e.	General User							
System Design:									
3	Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction -								
	from social media for CRM?								
✓	Yes	✓ 100%	80%	50%	30%	0%			
4	Do you think CRM concept can help identifying aspect for design?								
✓	Yes	✓ 100%	80%	50%	30%	0%			

5	Do you think process flow can help identifying weak points in current process?									
✓	Yes	✓ 100%	80%	50%	30%	0%				
6	Do you agree with Aspect design in this experiment?									
✓	Yes	✓ 100%	80%	50%	30%	0%				
7	Do you agree with measurement in this experiment by using word cnt and sentiment scoring?									
✓	Yes	✓ 100%	80%	50%	30%	0%				
8	Do you have any recommendations?									
	Do not have									
System Usage:										
9	Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction - which get from social media?									
✓	Yes	✓ 100%	80%	50%	30%	0%				
10	Can the results of data mining rule express the dissatisfaction which get from social media more concretely?									
✓	Yes	✓ 100%	80%	50%	30%	0%				
11	Which measurement do you prefer to use for analysis ?									
	a.	word cnt	b.	score cnt	✓	c.	both			
12	Do you prefer to use multidimensional sentiment cube for analysis ?									
✓	a.	Prefer and can analysis by yourself								
	b.	Prefer but cannot analysis by yourself								
	c.	Not prefer								
13	Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?									
✓	a.	Prefer								
	b.	Not prefer								
14	What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?									
	-									
System Validation:										
15	This result is for?									
✓	a.	Confirm your knowledge			✓100%	80%	50%	30%	0%	
✓	b.	New knowledge			✓100%	80%	50%	30%	0%	
16	Can these results help identifying the process which need to improve on dissatisfaction problem?									
✓	Yes	✓100%	80%	50%	30%	0%				
17	Can Multi-dimensional sentiment cube adapt to another domain?									
✓	a.	Can								
	b.	Cannot								
18	Advantage									
	This system can reveal hidden problems in real business with related features each other.									
19	Disdvantage									
	To manage Thai language may take a lot effort and spend high cost in real implementation.									
20	Recommendation									
	To update Thai language database (lexicon) more frequency to collect new or modern words.									

Interviewee#13 name: Mr.Jaturong Kerdrat (Other Domains with non-IT)
Position: Managing Director
Company: Getsmart Service
Contact: (+6683) 555-5391

1 Does multidimensional sentiment cube or sentiment analysis have already implemented in your organization?

✓ a. not implement
b. under study
c. never heart before

2 In your opinion, who is proper to use this system?

✓ a. domain expert
- ✓ a1. Insurance
- ✓ a2. CRM
- ✓ a3. Business Analysis
- a4. IT

✓ b. Management
✓ c. Strategic planner
d. User in organization
✓ e. General User

System Design:

3 Do you think the concept of sentiment analysis can help identifying the customers' dissatisfaction - from social media for CRM?

✓ Yes ✓100% 80% 50% 30% 0%

4 Do you think CRM concept can help identifying aspect for design?

✓ Yes ✓100% 80% 50% 30% 0%

5 Do you think process flow can help identifying weak points in current process?

✓ Yes ✓100% 80% 50% 30% 0%

6 Do you agree with Aspect design in this experiment?

✓ Yes ✓100% 80% 50% 30% 0%

7 Do you agree with measurement in this experiment by using word cnt and sentiment scoring?

✓ Yes ✓100% 80% 50% 30% 0%

8 Do you have any recommendations?

Do not have

System Usage:

9 Can the results of multidimensional sentiment cube express the relation among features of the dissatisfaction - which get from social media?

✓ Yes ✓100% 80% 50% 30% 0%

10 Can the results of data mining rule express the dissatisfaction which get from social media more concretely?

✓ Yes 100% ✓80% 50% 30% 0%

11 Which measurement do you prefer to use for analysis ?

a. word cnt b. score cnt ✓ c. both

12 Do you prefer to use multidimensional sentiment cube for analysis ?

✓ a. Prefer and can analysis by yourself
b. Prefer but cannot analysis by yourself
c. Not prefer

13	Do you prefer to use data mining rules for setting up the dependency of severity tasks to solve problems?								
✓	a.	Prefer							
	b.	Not prefer							
14	What different between multidimensional sentiment cube usage and data mining rules usage in your opinion?								
	-								
System Validation:									
15	This result is for?								
✓	a.	Confirm your knowledge		✓100%	80%	50%	30%	0%	
✓	b.	New knowledge		100%	✓80%	50%	30%	0%	
16	Can these results help identifying the process which need to improve on dissatisfaction problem?								
✓	Yes	✓100%	80%	50%	30%	0%			
17	Can Multi-dimensional sentiment cube adapt to another domain?								
✓	a.	Can							
	b.	Cannot							
18	Advantage								
	This system show the relationship of key information that is important to provide the reason and impact each other.								
	Cross analysis among features can express the analysis results in depth details.								
19	Disdvantage								
	This system concerns on loss data from user-generated contents which occur from language usage by slang words, -								
	fashion words or typo.								
20	Recommendation								
	In the real practice, it requires human to monitor real sentiment to confirm with system tracking in order to reduce -								
	mis-tracking from speaking language.								

Appendix B

Survey on satisfaction of life insurance services in Thailand

Respondents Profile:

1. Gender	a. Male	b. Female	c. unspecified
2. Age___________ years old			
3. Status	a. Single	b. Marriage	c. Widow
4. Monthly Income ___________ baht			
5. Occupation ___________			
6. Education level	a. Undergraduate	b. Bachelor's Degree	c. Master's Degree
	d. Doctor's Degree	e. Diploma or Certificate	
7. Do you have work experience or study in life insurance?	a. Yes b. No		

Life Insurance Characteristic of respondents *(If you do not have a life insurance policy. Answer only items 1 and 2 and skip to the next section on 'Life insurance attitude' without answering questions 3-10 in this topic.)*

1.	Do you have life insurance?	a. Have ___ policies	b. Do not have	
2.	Do you decide to buy life insurance for your family?		a. Yes	b. No

3.	Life Insurance company of your policy	a. AIA	b. Thai Life Insurance	c. Alianze Life Insurance
	d. SCB Life Assurance	e. Bangkok Life Insurance	f. Muang Thai Life Assurance	g. Others ________
4.	Type of Life Insurance which you hold	a. Endowment	b. Whole Life	c. Term
		d. Retirement	e. Investment	f. Others ________
5.	Period of coverage	a. Less than or equal 10 years	b. More than 10 years but less than 20 years	c. 20 years up
6.	Total of premium per year of all policies	a. Less than of equal 15,000 baht	b. 15,000 – 30,000 baht	c. 30,001 – 60,000 baht
		d. 60,001 – 100,000 bath	e. 100,001 – 200,000 baht	f. 200,000 baht up
7.	How many years have you been a life insurance customer?	….. years		
8.	Payment method	a. Monthly	b. Quarterly	c. Half year
		d. Yearly	e. One-time payment	
9.	Have you ever use service of life insurance company? Which service? (You can reply many answers)			☐ Never use service
	☐ Yes, I ever use service	a. General inquiry	b. Request make new life insurance policy	
		c. Request claim	d. Request make new rider such as H&S, Cancer protection	
		e. Request change status	f. Make policy cancellation	g. Others __________

10.	Which channel do you prefer to contact for getting service from life insurance company?			
		a. Agent	b. Broker	c. Bank
		d. Call center	e. Website	d. Government Organization

Attitude about Life Insurance

1.	Do you think life insurance is important?	a. Very important	b. Important	c. Not important
2.	How do you know the criticism about the service of life insurance?	a. Experienced	b. No direct experience. But realize from social media	c. No direct experience. And do not know any details
3.	When you experience an unsatisfied service, you will…	a. Negligent	b. Do not return to get service from that company	c. Spread the word to others
4.	From your experience about service in life insurance. What is your sentiment?	a. Positive	b. Negative	c. Neutral
5.	Do you believe in criticism which you've heard from social media such as Facebook, Line, Twitter, websites from celebrity interviews?	a. Absolutely believe	b. Quite believe	c. Do not believe
6.	In your opinion, why some person do not want to buy life insurance? (checkbox)	Do not want to pay because it does not necessary.	Feel annoyed by agents who offer insurance plans.	Strong health
	Life insurance companies are more beneficial than provide benefits.	Life insurance company and agent cheated	Another investment. More profitable	Not sure to claim success.
	There is insurance in other kind on hand	No one is worried.	Happy with social security	

Part 1. Evaluate knowledge about life insurance.

	Sentiment Level	1 Less	2	3 More	4	5 Most
1	Do you think how much you have any knowledge about life insurance?					
2	Before making decision to make a life insurance policy, do you think you understand the contents and exceptions in the policy?					
3	After receiving the life insurance policy, do you think you understand the contents and exceptions in the policy?					
4	Do you need more study about the contents and exceptions inside the policy than the agent explanations under limited time or not?					
5	Do you think that there are some terms or exceptions of the policy that you think you have misunderstood?					
6	Do you understand the meaning of the sum assurance, the contractual reimbursement, and the dividend on the life insurance?					
7	Do you understand about "Free look period" which is to cancel the life insurance policy within 15 days from the date the insured received the policy?					
8	Do you know? If you do not reveal your health history exactly, there is a risk that you will not be able to claim benefits or lead to policy cancellation.					
9	Do you understand? At the expiration period of the policy. Total sum assure of life insurance policy will be fully refunded under the contract. However the dividends may vary.					
10	Do you know? If you want to make policy cancellation or lapse, you can choose to change status of them instead such as reducing the sum assure or reducing the coverage period. This results in no premium payment. You do not have to cancel the policy, which will lose benefits.					

Part 2. Evaluate the service of policy offering, claim the policy, apply for change status, and make policy cancellation.

	Sentiment Level	1 Less	2	3 More	4	5 Most
1	Do you feel that you have been tempted to make your current policy? and you are not sure whether this policy meets your needs or not.					
2	Do you think that the insurance agent provides distortion or concealment of policy terms to customers?					
3	Please evaluate the disqualification of life insurance service nowadays.					
	a. Reliability and integrity in service					

	b. Product and service knowledge					
	c. Responsibility in service to customers					
	d. Continuous customer service throughout the policy period					
	e. Personality, service etiquette					
4	Have you heard or experienced of problems on the service process below?					
	a. Presentation and issuance of new policy process					
	b. Claim process when stay at hospital					
	c. Claim process to receive compensation					
	d. Policy change request process such as cancel policy or rider					
	e. Marketing and promotion process					
	f. Refund sum assured process. (when policy expires)					
5	Have you heard or experienced on any claim problems below?					
	a. Cannot contact agent or no service provider support					
	b. Do not understand in claim exceptions					
	c. The process of claim at the hospital is very delicate.					
	d. Long time to get compensation					
	e. Get compensation return less than expectation					
	f. Cannot claim from disease					
	g. It is inconvenience to make a reservation for medical expenses in advance.					
	h. Health record concealment problem					
6	From your experience, what is the reason why customers apply for changing policy status?					
	a. Request a change because of poor service					
	b. Request a change because the policy does not meet the requirements					
	c. Request a change due to demand or financial status has been changed					
7	From your experience, what is the reason why customers apply policy cancellation?					
	a. Make a cancellation due to poor service					
	b. Make a cancellation because the policy does not meet the requirements.					
	c. Make a cancellation due to demand or financial status has been changed					
	d. Make a cancellation due to want to change insurance company					
	e. Make a cancellation to take out insurance funds ahead of time					
8	When you get a dissatisfaction service, you choose to					
	a. Litigation company or agent					

	b. Post a topic in social media about the service that comes across					
	c. Spread the word to relatives and friends					
	d. Make a policy cancellation					
	e. Contact the management of the company in order to negotiate and solve problems.					
9.	Which of the following problems occur from a misunderstanding about the principle of life insurance between customers and service providers?					
	a. There are many policy conditions to understand in short term					
	b. Service providers such as insurance agents inform customers only the benefits of the policy which they offer for sales					
	c. Service provider does not check the customer's understanding before making a contract.					
	d. Economic situation is changed and some benefits such as dividends may impact.					
	e. Lack of service providers with sufficient knowledge					
10.	Please rate the misconception issues of life insurance service nowadays. (1 is less problem, 5 is a lot problems)					
	a. the scope of coverage					
	b. the conditions of life insurance coverage					
	c. the scope of coverage - disease					
	d. the duration of the disease					

Part 3. Evaluate the service improvement.

	Sentiment Level	**1 Less**	**2**	**3 More**	**4**	**5 Most**
1.	How do you give important to this service type?					
	a. Reliability and honesty in service					
	b. Product or service knowledge					
	c. Responsibility in service to customers					
	d. Continuous customer service throughout the policy period					
	e. Personality, service etiquette					
2.	Please rate the method which is possible to solve the dissatisfaction problem (1 = less help, 5 is most help)					
	a. Negotiations to compensate for the loss together					
	b. Offer an opportunity to change insurance type					
	c. Find a central authority organization to help negotiation such as the					

	Department of Insurance.					
	d. Discounts on annual premiums (In case the customer is correct)					
	e. Provide gifts such as free health check card					
3.	When customers want to make policy cancellation which cause to loss benefits, what do you think about these offers of status changing as the following?					
	a. Reduce sum assured by average the amount of premiums have already paid under the same protection period and customer can stop payment premium in the next year with active policy					
	b. Reduce period of coverage by average the amount of premiums have already paid under the same sum assured and customer can stop payment premium in the next year with active policy					
4.	What activities do you think will enhance your understanding of life insurance? To reduce misunderstanding between customers and service providers in the long run.					
	a. Product knowledge training to customers who hold that policy type					
	b. Training to provide understanding especially in highlighted problem.					
	c. Product knowledge and ethics training to agent					
	d. Send mail or message via social media to describe collateral and simple exceptions such as infographic, short-clips, cartoons					
	e. Make a contact customer for asking an understanding or provide any help on some problems, especially when the policy details are changed					
	f. Provide a simple questionnaire to customer for self-checking an understanding					
	g. Make a life insurance brochure placed by general place such as department store, BTS.					
	h. Organize a department to review customer opinion especially problem from social media.					
	i. Organize troubleshooting agencies for an individual support in the urgent case					
	j. Keep evaluate the customer satisfaction in each period					
5.	What do you think? If customer's attitude toward life insurance is better, it can make the growth of life insurance in Thailand?					
6.	Other recommendation for service improvement					

List of publications and presentations:

1. Sucharittham, N., Threeramunkong, T., Haruechaiyasak, C., Bao, H. T., & Hieu-Chi Dam, Data Mining for Life Insurance Knowledge Extraction: A Survey, *The Asian Conference on Information Systems (ACIS2013)*, Bangkok, Thailand, 2013.
2. Sucharittham, N., Haruechaiyasak, C., Theeramunkong, T., & Hieu-Chi Dam, Exploiting Linguistic Knowledge for Sentiment Classification in Online Opinions related to Life Insurance Topic in Thailand, *The Asian Conference on Information Systems (ACIS2014)*, Natrang, Vietnam, 2014.
3. Sucharittham N., Nguyen T., Theeramunkong T., Hieu-Chi Dam, Customers' Sentiment in Thai Life Insurance, *International Journal of Management, Business, and Economics*, 2020. (Accepted, 2020), 15 pages.
4. Sucharittham N., Nguyen T., Theeramunkong T., Hieu-Chi Dam, Customer Sentiment Knowledge Management: A case study in Thai life insurance, *International Journal of Management, Business, and Economics*, 2021. (Accepted, 2021), 14 pages.
5. Sucharittham, N., Haruechaiyasak, C., Hieu-Chi Dam, Threeramunkong, T., Multi-dimensional sentiment cube mining for process monitoring, *Walailak Journal of Science and Technology*, 2020 (Under review), 20 pages.
6. Sucharittham, N., Haruechaiyasak, C., Hieu-Chi Dam, Threeramunkong, T., A design of multi-dimensional sentiment cube on social customer relationship management, *Journal of Intelligent Information Systems*, 2021 (Under review), 25 pages.