Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 126 (2018) 986–995

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2018, 3-5 September 2018, Belgrade, Serbia

Obtaining Exhaustive Answer Set for Q&A-based Inquiry System using Customer Behavior and Service Function Modeling

Hironori Takeuchi[a,*], Satoshi Masuda[b], Kohtaroh Miyamoto[b], Shiki Akihara[c]

[a]*Musashi University, Toyotamakami 1-26-1, Nerima-ku Tokyo 176-8534, Japan*
[b]*IBM Research - Tokyo, Nihonbashi Hakozaki-cho 19-21, Chuou-ku Tokyo 103-8510, Japan*
[c]*Global Business Services, IBM Japan, Ltd. Nihonbashi Hakozaki-cho 19-21, Chuou-ku Tokyo 103-8510, Japan*

Abstract

When customers are interested in a service or intend to buy it, they sometimes have questions on that service. In this study, we considered an inquiry system in which customers ask questions on a specific service and obtain correct information on the service. For such an inquiry system, a question-answering (Q&A) technology is needed. Many programming modules for such a technology have been developed and can be easily used for system development. In many Q&A technologies, machine-learning techniques are involved, and we need to prepare training data consisting of pairs of an answer and assumed questions. For training-data preparation, an answer set for a service should be defined as the first step and the answer set should cover all the information on the service that customers may ask about. By using a customer-behavior model and introducing a service-function model, we propose a method of effectively collecting knowledge information for an answer set on a service. Through a case study, we show that we can collect exhaustive knowledge information for an answer set with our method compared to the case in which domain experts collect knowledge information in their own way. For an actual project, we also considered an actual inquiry-system-development project, with training data obtained with the proposed method, and showed that the system covers almost all the information on the service that customers may ask after a user test.

Keywords: Inquiry System, Answer Set, Customer Behavior Model, Service Function Model

1. Introduction

With the recent rapid expansion of artificial-intelligence (AI) technologies, many machine-learning-based programming modules such as text classification and image recognition, have been developed and made available as application programming interfaces (APIs). Developers do not have to worry about the details of the machine-learning algorithms but use the module function by just preparing the training data required for the machine-learning programming module. When developing an application for a service provider, developers need to develop training data for each application with the help of domain experts of the service provider. The output of a machine-learning-based

* Corresponding author. Tel.: +81-3-000-0000
E-mail address: h.takeuchi@cc.musashi.ac.jp

10.1016/j.procs.2018.08.033

system is probabilistic, and the output may change dramatically when the input slightly changes. Therefore, not only the application interface but also the output quality of the machine-learning component is important. The quality of output mainly depends on the training data, and training data should be developed carefully. However, there are few methods on training-data development that can be applied in a wide area.

In this study, we considered a situation we develop an inquiry system that involves customers asking questions about a specific service. The inquiry system automatically responds to an inquiry from a customer on behalf of the domain expert. Question and Answering (Q&A) is a technology of automatically answering a user's question[1 2] and is a core technology for an inquiry system. With a Q&A technology the query is first extracted from the question. After that, a set of answer candidates is found from the answer collection with this query. Finally, a relevance score is assigned to each candidate for the output. An inquiry system sends the input text to the engine implementing the Q&A technology and obtains a set of answer candidates. From the candidates, the system will show a pre-specified number of answers with high scores as output. There are implementations on the Q&A engine that are published as an API. In this paper we consider a Q&A approach based on a text-classification technology [1].

Through Q&A programming modules using machine-learning techniques, developers do not have to worry about such algorithms and only have to consider the training data when using the module as an API. In a text classification-based inquiry system, we first define an answer set $A = \{a_1, a_2, \ldots, a_n\}$ where n is the size of A, and collect questions assumed for each answer. The assumed questions of answer a_i are defined as $Q_i = \{q_{i1}, q_{i2}, \ldots, q_{im}\}$. Here, m is the variation size of the assumed questions (ex. $m = 10$). We prepare an A and assumed questions $Q_1, \ldots, Q_n$ as training data and apply the data to the machine-learning module in the text-classification engine. After the training phase, the system can obtain a question text and output the answers with high categorization scores.

In this situation, we consider a method of effectively preparing training data. The inquiry system responds to customers on behalf of the experts of the service provider, so the quality of answers that the system outputs is important for customer satisfaction. First, the inquiry system should answer all questions from customers. This means that the developer should list all the business knowledge for the A through cooperation with the service provider. Second, the system should answer questions from customers correctly. This means that a wide variety of assumed questions Q_i should be identified for each a_i. In a test of the developed inquiry system, the system may sometimes not correctly respond to the test question. If the expected answer exists in the training data, the developer has to add new assumed questions for it so that the system will output the answer properly. If the expected answer does not exist in the training data, it should be defined first because it is not included in the current A. After that, the assumed questions for the answer should be collected. In this case, the service provider has to be involved to define the expected answer. Therefore, to reduce the cost for thoroughly modifying training data when an error occurs in a test, we need to carefully define an A in the design and implementation phases so that it includes all service knowledge that the customers want to know.

We propose a method of obtaining an exhaustive A for the training data based on the modeling of customer behavior and service functions so that an inquiry system can correctly answer any questions that customers may ask. We validated the effectiveness of our method through case studies. The rest of this paper is as follows. In Section 2, we describe related work. In Section 3, we describe the proposed method for obtaining an answer set for an inquiry system. In Section 4, we discuss the effectiveness of the proposed method through case studies. After the discussion in Section 5, we conclude the paper in Section 6.

2. Related Work

With the rapid expansion of machine-learning technologies, many machine-learning algorithms can be made available as APIs. Developers do not understand these algorithms but use them through APIs and construct many business-to-consumer (B2C) or business-to-Business (B2B) applications on various devices. For agile development of machine-learning technology-based applications, a reusable data model and application architecture have been introduced so that developers can construct an application and validate it in a real environment based on the well-known Agile

[1] For example：https://www.ibm.com/smarterplanet/us/en/ibmwatson /developercloud/doc/nl-classifier/index.shtml

methodology[3]. Schreiber et al.[4] introduced a monitoring system to find the best-performing machine-learning API combination to reach robust output quality for the different domains. However, there is a lack of methodologies for preparing data resources, such as training data, and in many cases developing training data for a machine-learning-based application depends on practitioner skill and experience.

There have been studies on a practical Q&A system. In these studies, the data source required for such a system was discussed. In the open-domain question-answering area, knowledge bases with over one-hundred million triples constructed from structured public data sources, such as Freebase and Probase, are used for the system[5 6]. On such answer-data sources for open questions, a quality assessment has been conducted[7]. In this assessment, data coverage is not discussed though availability, consistency and trustworthiness are considered as quality metrics. In the closed-domain question-answering area, domain specific concepts are defined in advance, and such concepts are used for the answer candidates[8]. Also, automatically collecting answer is done through the domain community or conversational dialog with experts[9 10]. In these studies, though the quality of answer contents is assessed, there have been no methodologies introduced to assess whether the obtained answer set fully covers the domain.

3. Obtaining Answer-set using Customer-behavior and Service-function Modeling

In this section, we explain our method for obtaining an A based on the modeling of customer behavior and service functions.

3.1. Method Overview

An inquiry system should answer all questions from customers; therefore, all the business knowledge for an A should be listed through cooperation with the service provider. Also, the system should answer questions from customers correctly, therefore, a wide variety of assumed questions Q_i should be identified for each a_i.

Software-quality characteristics are considered in ISO 9126[11 12], which identifies six main quality characteristics, i.e., Functionality, Reliability, Usability, Efficiency, Maintainability, and Portability. Functionality is the essential purpose of any service and is divided into sub-characteristics, i.e., Suitability, Accurateness, Interoperability, Compliance, and Security. We argue that the following two sub-characteristics are important for an inquiry system.

- Suitability: appropriateness to specify the functions of the system
- Accurateness: correctness of the functions

Listing all the business knowledge and a wide variety of assumed questions are required for satisfying Suitability and Accurateness respectively. As described in the introduction, in-suitability of the training data will have a huge impact on system development; therefore, our proposed method is used for obtaining an A covering all the knowledge of a service to satisfy Suitability metrics.

An inquiry system is considered a human-machine interaction systems. Through research on user-behavior analysis on practical human-machine interaction systems, it was found that users will ask questions to a system in their various situation and such questions are sometimes out of the scope of the service[13]. As a result, there are four cases regarding the relation between customers questions and the inquiry-system output.

(A) A customer asks a question on a service and the expected answer is defined.

(B) A customer asks a question on a service but the expected answer is not defined.

(C) A customer asks a question not related to a service, and the expected answer is not defined.

(D) A customer does not ask a question on a topic that is defined in the system.

Our proposed method for obtaining exhaustive A reduces the the occurrence of case (B) because case (C) and (D) can be detected only through user tests.

To list all answers for potential questions from customer on a service, we need to consider what type of information customers will need and what service elements customers will refer to in their questions. Service elements are consisting the service. For example, commission charges, user password, and term of balance inquiry are service elements for an online banking service. Customer need various types of information on a service element, and they will depend on each customer's context and element features. Therefore we introduce two models to list all knowledge on a service for a high-coverage answer set.

- Model for relations between customers' information needs and their context
- Model for elements consisting of a service

Following sections, we describe these models and our method for obtaining an A.

3.2. Knowledge Types for each Customer-decision Process

To consider the relation between customers' information needs and their context, we first consider the model of customer behavior. Customer behavior involves many personal and situational variables, and many customer-behavior models have been proposed[14]. These models are abstract representations of the customer-decision process and customer behavior. In this study, we focused on a customer-decision-process model called the Engel-Blackwell-Miniard (EBM) model[15]. In this model, the process begins with the stimulation of a need. After identifying the need, the customer searches for information on the product, and a set of preferred alternatives will be generated. The customer will evaluate the alternatives and a purchase will be made based on the chosen product. Post-purchase evaluation will affect future decision making. These steps are shown in Figure1. Based on this EBM model, we

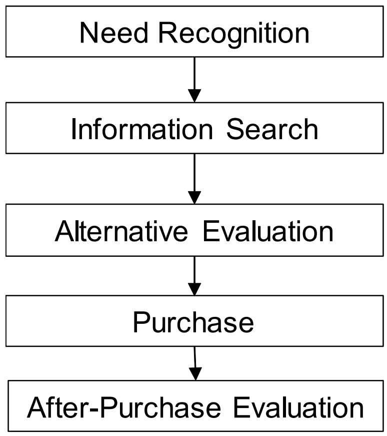

Fig. 1. Customer-decision-process model

considered three buying stages of a purchase as a customer's context.

1. Recognition stage: Need recognition
2. Emotion stage: Information search and alternative evaluation
3. Action stage: Purchase and after-purchase evaluation

Second, we consider customers' information needs. The type of information that the customer wants to know is considered a question type in the Q&A technology. Some studied were focusdc on a social Q&A site and investigated what types of information are requested in the Internet community. Kim et al.[16] considered that questions in a Q&A site are divided into three groups:

- Information: finding specific facts or understanding phenomenon
- Suggestion: seeking advice or recommendations
- Opinion: surveying other people's thoughts or initiating discussions

Rodrigues and Milic-Frayling[17] defined the following two groups in addition to the above three:

- Chatting: communicate with other users at a personal level
- Entertainment: posting trivia or puzzles for the community to solve

Based on these definitions, we define categories for questions required for an inquiry system. In a situation in which a customer wants to obtain knowledge on a specific service through an inquiry system, questions are divided into information and suggestions, and the customer wants to know the various types of information. Therefore, we define the following four categories for questions as knowledge types of a service required for an A.

1. Fact: explaining facts or definitions on a service element
2. Procedure: explaining how to use a function in the service
3. Reason: explaining the customer's situation when using a service function and having problems
4. Suggestion: describing the suggestion or recommendation of a service function

We organize the information on a service with these types and call the organized information "knowledge information". Each entry in the knowledge information corresponds to the answer description a_i in an A. Therefore, an A can be obtained by identifying elements comprising the service and defining each type of information for the elements.

In recent studies referred above, knowledge types for a question were fixed and static. As described in 3.1, however, customers will be interested in a different aspect on the service based on their situation. For example, when recognizing a service, a user wants to know what the service is. He wants to know the details or learn about how the service is used when he is interested in buying the service. After purchase, he sometimes needs a solution when having problems with the service. This means that the knowledge types for questions considered in an inquiry system will depend on the target users of the system.

We now introduce the relation model between knowledge types (information needs) and a customer's buying stages (context). Table 1 shows the relation between knowledge information and buying stages. In Table 1, the degree of

Table 1. Relation between knowledge information and buying stages (H:high, M:medium, L:low)

	Recognition Stage	Emotion Stage	Action Stage
Fact	H	H	M
Procedure	M	H	H
Reason	L	L	H
Suggestion	H	M	L

information requested by customers is represented as high, medium, or low. When the customer is becoming familiar with the service and is going to buy it, he will not need Fact information of the service but will need Procedure or Suggestion information. After purchasing the service, the need for Suggestion information decreases but the need for Reason information will increase. By using this model, after defining the type of customer of the inquiry system, we can determine which types of knowledge information on the service should be collected for the system. For example, Fact, Procedure, and Suggestion information should be collected for the inquiry system in which customers in the Recognition or Emotion stage will ask a question,

3.3. Service-function Analysis and Obtaining Answer-set

We now consider our proposed method of identifying all the elements comprising a service and obtaining an A using the relation model introduced in the previous section.

First, we introduce a service-function model. In this model, a service consists of the following three elements.

- Service function: the function the service provides
- Variable attribute: the attribute, the value of which the customer can change

- Invariable attribute: the attribute, the value of which the customer cannot change

Figure 2 represents the relations between these elements. By using this model, we can identify the service elements.

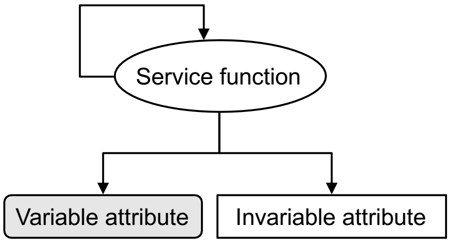

Fig. 2. Service-function model

Figure 3 shows the steps of service-function analysis. We start with a top-level service function and list its variable

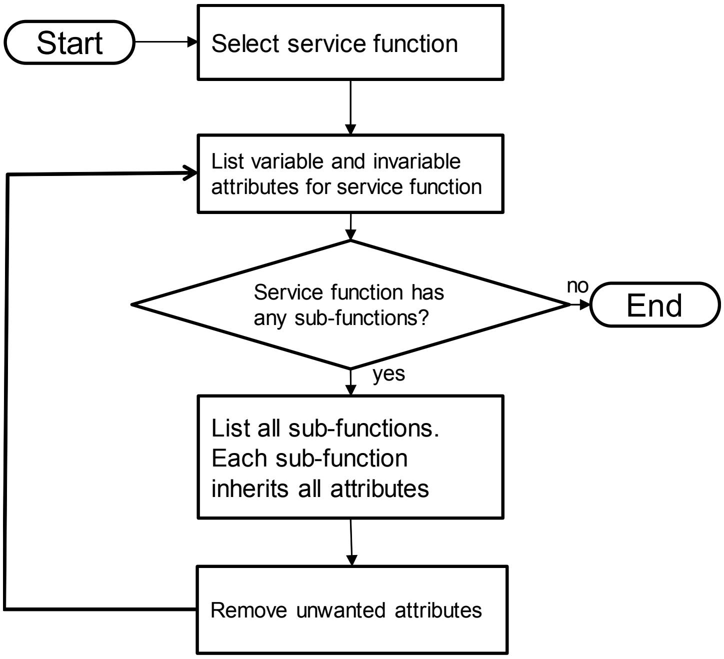

Fig. 3. Steps of service-function analysis

and invariable attributes. Next, we list sub-functions under the service function. The sub-functions inherit all the attributes of the parent function. After assessing each sub-function, we remove non-required attributes and add the sub-function-specific attributes. We continue these steps and finalize the analysis when the function does not have any sub-functions. Figure 4 shows the results of the service-function analysis in which an online-banking service is considered.

By using the service elements identified from the service-function analysis, we organize the knowledge information. The knowledge information collected for each service element depends on the element type. For example, we can only define Fact information for invariable attributes because we cannot change or suggest anything on the attribute. For each element in the service-function model, we introduce a model of knowledge types required for each element in the service-function model, as shown in Table 2.

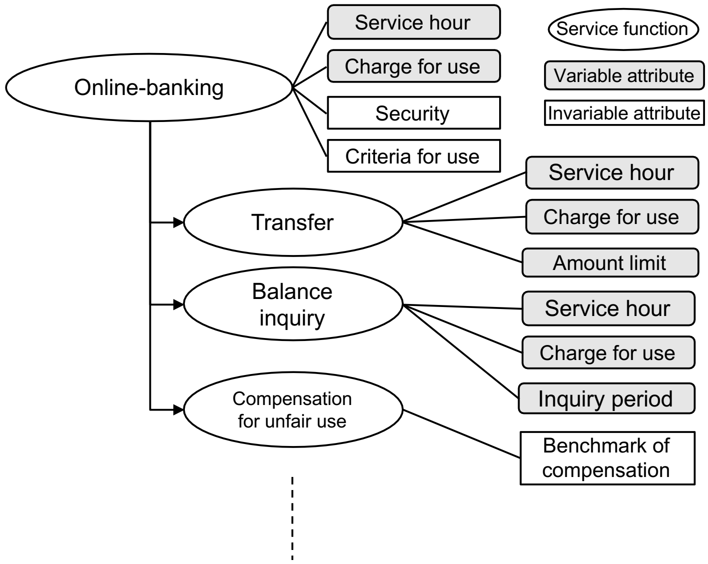

Fig. 4. Service-function-analysis results

Table 2. Relation between service elements and knowledge information (Y: defined, N: not defined)

	Fact	Procedure	Reason	Suggestion
Service function	Y	Y	Y	Y
Variable attribute	Y	Y	Y	N
Invariable attribute	Y	N	N	N

Now, we can organize the knowledge information for the service elements identified from the service-function analysis by using the relations described in Tables 1 and 2 after defining the customer-buying stages for the inquiry system,

The result of organized knowledge information is used as A in the training data for the Q&A module. Table 3 shows examples of organized knowledge information on an online-banking service.

4. Case Studies

4.1. Comparison of methods of obtaining Answer-set

We evaluated the effectiveness of our proposed method for a banking service. In this experiment, we simulated developing an inquiry system for an overseas remittance service and the target customers of the system will be in the Recognition or Emotion stage. For such an inquiry system, two domain experts obtained the As.

One system engineer without any domain knowledge developed an A based on our proposed method. As a result, 5 service functions, 0 variable attributes, and 14 invariable attributes were identified and an A consisting of 29 pieces of knowledge information was obtained. In contrast, another engineer with banking-domain knowledge obtained an A in his own way without any procedures or guides. As a result, an A consisting of 14 pieces of knowledge information was collected. We could collect much more knowledge information on the service with our proposed method. From this study, we can expect to obtain an exhaustive A without depending on a developer's skill or experience.

Table 3. Examples of organized knowledge information (answer set)

Topic	Fact	Procedure	Suggestion
Online banking	Online banking provides banking services through your PC or smart-phone.	You can apply for the online banking service when you open your bank account.	Online banking is very useful because you can use the banking service without visiting the branch office.
Term of balance inquiry	This is the term in which you can check your balance through your PC or smart-phone.	You can change the term up to 2 years through the setting screen.	N/A
Benchmark of compensation for unfair use	Customer whose account is used by someone else can ask for compensation.	N/A	N/A

4.2. Testing in Inquiry-system-development Project

We applied our method to a real inquiry-system-development project for a service provider. The system automatically responds to questions on a service from customers who were interested in a service. The client interface was located at the branch office of.

Before actual customers officially used the inquiry system, we conducted a system-user test. Because the system responds to customers by synthesized voice, this inquiry system will present one answer as an output. In the user test, we validated whether the system presented the answers that the subject-customers expected.

We determined three cases from the test results.

(i) Expected results were presented

(ii) Expected results were not presented, though the correct answer was included in A

(iii) The customer required the information that should have been included in A. As a result, the system could not answer the question correctly.

(iv) The customer required the information not directly related to the service and not defined explicitly by the service provider. As a result, the system could not answer the question correctly.

In the cases in which expected results were not presented, the system did not satisfy Accurateness in case (ii) and Suitability in cases (iii) and (iv).

We conducted a user test involving customers and updated the training data by adding new answers and enhancing assumed questions. We then made the system available to actual customers. By analyzing the log data in the system, we collected the questions from actual customers and answers from the system and evaluated whether the system correctly responded to the customers' questions. We conducted two evaluations of the actual usage data and updated the training data after each evaluation. Our method obtained an A satisfying Suitability. Therefore, we assessed how much the number of unexpected results decreased through the test and evaluation of the real usage data. Figure 5 shows the analysis results of the answers from the inquiry system. In Figure 5, Set 1 represents the user test results and Sets 2 and 3 represent the evaluation results of actual customer usage. We found that the number of cases not satisfying Suitability decreased after the user test.

5. Discussion

In a system-development project, the system will need to be re-worked when a defect is found. Therefore, it is important to find a defect requiring a large amount of re-working in the early development phase. When the training

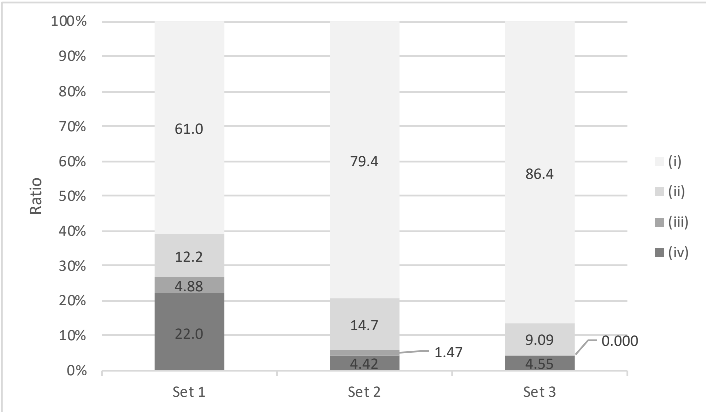

Fig. 5. Causal analysis of test results

data in an inquiry system do not satisfy Accurateness (case (ii) in the case study of the system-development project), the new assumed questions should be added for the expected answer so that the system will respond correctly. In this situation, the assumed questions enhanced by the developer will be checked by the service provider. However, when the training data do not satisfy Suitability (cases (iii) or (iv)), new knowledge information should be first defined as an answer by the service provider. After that, the developer will collect the assumed questions for the answer, and these questions will be checked by the service provider. In this situation, the cost required for the re-working will increase.

From the viewpoint of system-develop management, it is important to determine whether the training data satisfy Suitability in the requirement or design phase. In such a check, both the developer and service provider need to assess whether the obtained A covers all the information on the service and need to agree on its sufficiency. By using the organized knowledge information with our method, it is easy for all stakeholders to build a consensus on the suitability of the training data. It is also important to estimate the cost of developing the training data before proposing a system-development project. We can predict the size of the knowledge information by analyzing a part of the service and accurately estimate the training-data-development cost before proposing the project.

In this study, we considered the training data for the machine-learning module in an inquiry system. The training data consist of pairs of an answer and assumed questions. We proposed a method of obtaining an A from scratch by modeling customer-behavior and service functions. In many services, the service provider collects questions from customers at the customer support center, organizes them as frequently asked questions (FAQs), and publishes them on the web site so that customers can find solution themselves. Such FAQ data are mainly for customers who already use the service and do not contain questions that are related to the core service function but are not asked frequently. We cannot directly use the FAQ data as the training data but use them as a seed to organize knowledge information. By applying natural-language-processing technologies to both FAQ data and business documents (e.g. manuals, sales guides), we can expect to semi-automatically organize knowledge information for an A in training data. We consider this as future work.

6. Conclusion

In this study, we considered an inquiry system in which customers ask questions on a specific service and obtain correct information on that service. In the inquiry system, a Q&A technology has an important role, and many programming module for Q&A technology have been made available as APIs. When using a machine-learning module such as categorization-based Q&A technology, we need to prepare training data that contain pairs of an answer and assumed questions. We showed that the quality of the inquiry system depends on the quality of the training data, and the training data should satisfy Suitability and Accurateness. We should first obtain an exhaustive answer set in training data to satisfy Suitability.

To solve such a challenge in practical system development regarding training data, we proposed a method of obtaining an answer set by organizing knowledge information on a service required for an inquiry system. We introduced the buying stages based on a customer-behavior model and defined the knowledge information types for each buying stage. We also defined a service-function model and introduced steps in collecting service elements using this model. By using collected service elements and the relations between the required knowledge-information types, we organized knowledge information for an answer set in training data.

Through a case study, we showed that we could collect service elements exhaustively with the proposed method compared to the case in which domain experts collect knowledge information in their own way. By applying the proposed method to an inquiry-system-development project, we showed that we can collect almost all answers that customers may ask and avoid cases in which the training data do not satisfy Suitability and the developer and service provider should define a new answer for a question that the system cannot respond to correctly.

In many services, the service provider collects questions at the call center. For future work, we will modify the proposed method to organize knowledge information for an inquiry system by using FAQs and develop a semi-automatic method for organizing knowledge information.

References

1. Jurafsky, D., Martin, J.H.. *Speech and Language Processing*. Prentice Hall; 2008.
2. Manning, C.D., Raghavan, P., Schuetze, H.. *Introduction to Information Retrieval*. Cambridge University Press; 2008.
3. Schleier-Smith, J.. An architecture for agile machine learning in real-time applications. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2015, p. 2059–2068.
4. Schreiber, M., Kraft, B., Zündorf, A.. Cost-efficient quality assurance of natural language processing tools through continuous monitoring with continuous integration. In: *Proceedings of the 3rd International Workshop on Software Engineering Research and Industrial Practice*. 2016, p. 46–52.
5. Fader, A., Zettlemoyer, L., Etzioni, O.. Open question answering over curated and extracted knowledge bases. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2014, p. 1156–1165.
6. Yao, X., Durme, B.V.. Information extraction over structured data: Question answering with freebase. In: *Proceedings of the 2014 Annual Conference of Association for Computational Linguistics (ACL)*. 2014, p. 956–966.
7. Thakkar, H., Endris, K.M., Gimenez-Garcia, J.M., Debattista, J., Lange, C., Auer, S.. Are linked datasets fit for open-domain question answering? a quality assessment. In: *Proceedings of the 6th International Conference of Web Intelligence, Mining and Semantics*. 2016, p. 19:1–19:12.
8. Damiano, E., Spinelli, R., Esposito, M., Pietro, G.D.. Toward a framework for closed-domain question answering in italian. In: *Proceedings of the 6th IEEE International Conference on Signal-Image Technology & Internet-Based Systems*. 2016, p. 604–611.
9. Hixon, B., Clark, P., Hajisjirzi, H.. Learning knowledge graphs for question answering through conversational dialog. In: *Proceedings of the 2015 Annual Conference of the North American Chapter of the ACL*. 2015, p. 851–861.
10. Toba, H., Ming, Z.Y., Adriani, M., Chua, T.S.. Discovering high quality answers in community question answering archives using a hierarchy of classifiers. *Information Science* 2014;**261**:101–115.
11. ISO/IEC TR 9126: Software engineering - product quality. 19-12-2000, .
12. Bhatti, S.N.. Why quality?: Iso 9126 software quality metrics (functionality) support by uml suite. *SIGSOFT Software Engineering Notes* 2005;**30**(2):1–5.
13. Suchman, L.A.. *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press; 1987.
14. Teo, T.S.H., Yeong, Y.D.. Assessing the consumer decision process in the digital marketplace. *The International Journal of Management Science* 2003;**Omega**(31):349–363.
15. Rice, C.. *Understanding Customers*. Routledge; 1997.
16. Kim, S., Oh, J.S., Oh, S.. Best answer selection criteria in a social q& a site from the user-oriented relevance perspective. In: *American Society for Information Science and Technology Annual Meeting*. 2007, p. 1–15.
17. Rodrigues, E.M., Milic-Frayling, N.. Socializing or knowledge sharing? chracterizing social intent in community question answering. In: *Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM)*. 2009, p. 1127–1136.