Selection of Requirement Elicitation Techniques Using Laddering

Ja-Hee Kim
Seoul National University of Science and Technology, Seoul, Republic of Korea
e-mail: jahee@seoultech.ac.kr

Seung Mo Ham
Daebo Communication and System Corporation, Seoul, Republic of Korea

and Hwan-Ju Cha
SK Holdings, Seoul, Republic of Korea

Kwang Kook Kim
KOSCOM Corporation, Seoul, Republic of Korea

***Abstract*—Requirement elicitation is critical to the success of a project, but it is not known which techniques can identify the maximum requirement of the customer. Therefore, many researchers have proposed various requirement elicitation techniques and studied the pros and cons of these techniques. According to the results of these studies, there is no requirement elicitation technique that is optimal regardless of its environments, such as attributes of the project, elicitors, and stakeholders. In this paper, we analyze the cognitive structure of IT professionals based on the factors that actually affect the use of the requirement elicitation techniques using laddering. We also use various statistical techniques to examine the relationship between the attributes of the actual project and the requirement elicitation techniques.**

***Index Terms*—requirement elicitation techniques; laddering;decision tree;logistic regression model**

I. Introduction

That requirements are SMART (specific, measurable, attainable, realizable, time bounded) is regarded as an important factor that determines the success or failure of a project for improving the quality of a software [1][2][3]. Recently, researchers have developed dozens of different requirement elicitation technique (RETs) to cover the increasingly widespread use of software and a variety of project attributes. However, elicitors inundated with these techniques have a new agenda for what RETs to choose. Therefore, many researchers have conducted studies on how to choose RETs for a given project. In early days, they tended to propose frameworks for helping decision-making considering the advantages and disadvantages of RETs from the viewpoint of the techniques. However, the studies are gradually evolving from the project point of view, by gathering the factors that influence the decision-making of the requirements and by providing the criteria for selecting the RETs for a given project through theoretical or empirical analysis.

One thing to note when choosing RETs based on the attributes of a project and the stakeholders is that the types of attributes are affected by the social and cultural environment. Therefore, in this paper, we propose a method to gather the attributes to the framework rather than to provide a set of attributes for selecting RETs. In other words, we propose a framework to identify the attributes influencing the selection of RETs using laddering, which is one of the RETs, and to analyze the relationship between the RETs applied to the project and the attributes of the projects empirically. A study using a similar framework, although it was not a study of how to crawl the requirements, was conducted to analyze the criteria for product selection. This study used laddering and artificial neural networks to analyze the selection criteria of products such as golf clubs [4]. In this study, we analyze binomial logistic regression and decision trees for discriminant analysis.

This paper is composed as follows. In Section 2, we introduce the existing approach for selecting RETs and laddering, which is the main analysis tool in this study. Section 3 describes how the study was conducted, and we analyze the results in Section 4. Section 5 concludes this paper.

II. Related Works

A. RETs

Though it is essential to communicate stakeholder's needs to the elicitor through an intensive exchange of opinions, it is difficult to gather SMART requirements, because even the stakeholder cannot know exactly what he or she wants. For this reason, a number of RETs have been proposed, but there is no elicitation technique to solve all the problems like a silver bullet. That is, there are advantages and disadvantages to each technology, and so it is necessary to use a combination of several RETs after considering the situation. Thus, various frameworks have been proposed to help analysts find optimal RETs. The frameworks utilize the pros and cons of RETs [5][6], decide based on attributes affecting the selection of RETs [7][8][9], or statistically analyze real project data [10].

Early frameworks analyzed the strengths and weaknesses of RETs, designed a structural interview based on the results, and selected appropriate RETs through the interview [5]. This kind of framework should utilize surveys containing the pros and cons of each RETs [11][12]. This kind of paper classifies RETs into classic/tradition, cognitive, collaborative, contextual, conversational, observational, analytic, and synthetic, according to the communication method, and compares the advantages and disadvantages accordingly [7][11][13]. In this paper, we select RETs evenly for each category; that is, we select nine techniques including interview, survey, workshops, brainstorming,

978-1-5386-1107-4/17/$31.00 ©2017 IEEE

observations, document analysis, prototype, storyboard and JAD/RAD.

Other frameworks list the factors that influence the choice of RETs and recommend appropriate RETs based on the evaluation of the attributes of a given project [8][9][10]. To use this framework, we need to gather the attributes that affect the selection of RETs. Answar and Razali collected the attributes affecting the decision of the RETs through the existing research and an expert questionnaire [14]. These attributes were categorized through content analysis into four categories: stakeholder characteristica, requirements sources, technique feature and project environment. After that, elicited various attributes affecting the choice of RETs were classified into elicitor, stakeholder, project, problem domain, and elicitation process [8][10]. Additionally, because of the different preferences for RETs, depending on the country, culture, and age [7][15], it was also necessary to provide a framework for extracting those attributes but not just identifying them. In addition to research to identify as many attributes as possible for research on attributes [8][10], researches were conducted to suggest strategies for selecting RETs based on the core values of attributes [9].

B. Laddering

The means-end chain theory says that the attributes of a product or service to consider when you make a decision are to obtain a consequence, which is a means to achieve the ultimate value [16]. The theory generally uses the laddering technique for a semi-structural interview and survey and visualizes the results with a hierarchical value map (HVM). Soft laddering is a semi-structured in-depth interview technique that identifies attributes and repeats the formal questions to grasp the connection from the attributes to the participants' personal values via consequences. An attribute is the physical, sensory, abstract components or observable features that determine a decision. Consequences are more abstract than attributes and are positive or negative cognitive benefits that participants feel. Finally, personal values refer to the ultimate goal pursued in participants' lives [17]. To supplement the quantitative analysis of soft laddering, additional surveys are conducted, which is called hard laddering. Hard laddering makes codes summarizing the attributes, consequences, and values derived from the soft laddering. The next task is a questionnaire survey on the relevance of each item on attribute-consequence, consequence-value, and the relevance is summarized as an implied matrix. Finally, an HVM is created by analyzing the connection frequency of the step items shown in the implied matrix. The HVM is a kind of graph that shows the interconnections between attributes, consequences, and values. However, it is too complicated to determine the dominant cognitive structure of respondents if we show all the relationships between factors. Therefore, to more easily grasp the whole cognitive structure, the cutoff is set [18].

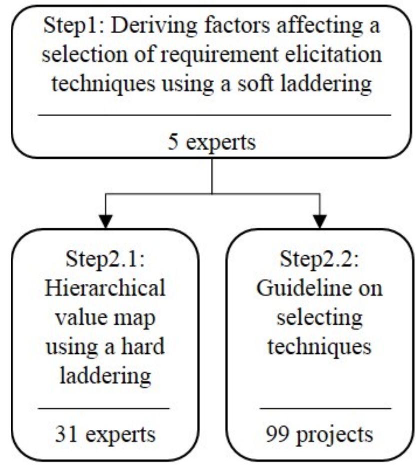

Fig. 1. Experimental process

III. Experimental Process

Fig. 1 explains our experimental process. In this paper, we first conduct an in-depth interview of Korean IT professionals about how and why they selected specific RETs for the projects that they had conducted. The result of the interview helps us model their value system. In the second step, we conducted an empirical analysis to give insight as to what project characteristics are required when seeking specific value in the project. Additionally, we examine the relationship between the attributes that influenced their decisions and the RETs that they actually used for their project. The detailed attributes and statical results may vary from country to country, but the framework that we suggest in this paper is expected to be common to all countries.

The first step is an in-depth interview with five experts having a Information Technology Professional Engineer qualification that can be acquired after at least nine years of IT experience using a soft laddering technique. In other words, we ask them to select one project that they have experienced and then to choose some RETs for the project. At the time, we ask why they select or do not select an RET. The following conversation illustrates the soft laddering.

Conv1
- Interviewer: Why did you determine to use the document analysis techniques, not the brainstorming technique, in the project?
- Interviewee: Because the document is updated well.

Conv2
- Interviewer: Why is it important that the document reflects the current system well?
- Interviewee: I believe a well-updated document can reduce the risk that the project may meet.

Conv3
- Interviewer: Why is reduced risk important for the project?
- Interviewee: Because it affects the quality of the project.

Conv4
- Interviewer: I got it. Do you have any other attribute that can decrease the risk?

- Interviewee: According to my experience, a project of which the user validated the requirements tends to have fewer risks.

From this dialogue, we can derive two attributes, one consequence, and one value. Conv1 shows that the attribute can include “how well the document reflects the current system”, and Conv2 means that the consequence of the attribute is “reduced risk”. We also find that the ultimate value that the interviewee wants to achieve from this attribute and the consequence is quality using Conv3. To find additional attributes, the interviewer uses Conv4 and finds an attribute called “user’s requirement validation”. Through such interviews, we derived 11 attributes, 8 consequences, and 2 values.

Step 2.1 is the quantitative analysis of the relationship among attribute-consequence-value (ACV) derived using soft laddering. We build an HVM, which summarizes the respondents’ cognitive structure. To do this, we design questionnaire using ACV and receive questionnaire responses from 31 experts with experience with IT projects. The demographic distribution of the respondents was 3 customer (9.6%), 19 developers (61.3%), 5 PMOs (16.1%) and 4 others (12.9%). For the hard laddering, we use the association pattern technique (APT) to construct an implication matrix for attribute-consequence and consequence-value. In other words, each column and row of the attribute-consequence implication matrix is assigned to consequence and attributes, respectively. Then, the respondent selects one to three consequences that he or she thinks the most important for each attribute. Similarly, the respondent fills the consequence-value matrix.

Step 2.2 is an empirical analysis of which RETs are used for each real project having a certain attribute. For this analysis, 98 projects are evaluated on a 7-point scale in terms of the attributes derived from Step 1. For example, let’s suppose that a respondent assumes a project that builds Company A’s homepage and that “updated document” has been derived from an attribute from the soft laddering. He gives seven points if all documents reflect the current system well, but 1 points if there is no document. He investigates all attributes in this way. Finally, he of she selects the RET used in the project. In this paper, we investigate nine techniques: interviews, surveys, workshops, brainstorming, observations, document analysis, prototype, storyboard, and JAD. To apply the results in the RET recommendation system, a logistic regression model and a decision tree model were generated for each attribute and “the use of RET” was used as an independent variable and a dependent variable, respectively. These analyses are conducted using SPSS.

IV. Analysis and Results

This section analyzes the results obtained through the procedure in Section III. In the first subsection, we visualize the relationship among ACV using HVM. In the second subsection, we examine the statistical results of step 2.2. The final subsection illustrates how to determine whether an RET is selected for a given project.

A. Relation among ACV

To examine all ACV derived from the soft laddering analysis, we record all ladders of each respondent, classify them into three levels of ACV, and summarize them in a summary code. In this study, we found 11 attributes, 8 consequences, and 2 values.

1) Attributes: Even without the knowledge about the existing research, the respondents derived similar attributes to the previous studies. These attributes are classified into elicitor (a6, a7, a8), informant (a4, a5, a9, a11), problem domain (a1, a3) and elicitation process (a2, a10)[8].

a1 clear project vision
a2 well-defined requirements process
a3 updated document
a4 the number of related organizations
a5 amount of requirements
a6 competent PM
a7 domain knowledge
a8 technical skill
a9 stakeholder engagement
a10 delivery period
a11 validation by users

2) Consequences: The benefits that we expect from good attributes are the following eight consequences:

c01 well-performed design stage
c02 easy tracking, managing, and implementing requirements
c03 easy communication with stakeholders
c04 convenience of scheduling
c05 low-change requirement
c06 reduced risk factors
c07 reduced overhead
c08 possibility of success

3) Values: When choosing a RET, the ultimate values of the experts finally converge on the quality of the project deliverables and the cost of carrying out the project at a reasonable cost.

v01 quality
v02 cost

The next step is to build a HVM using quantitative APT analysis. When connecting all the relationships in the HVM, it is necessary to determine the cut-off level; otherwise, the line is too complex making it difficult to grasp the main associations. If the cut-off is too big, some important relationship is ignored, and if the number is too small, you cannot know which connection is important. In this study, the numbers below 7 are cut-off because more than 70% of the linkages should be displayed in the previous study.

An HVM shows which attribute should be satisfied to achieve a consequence and which values can be met if the consequence is true. For example, if the vision of the project is clear (a1), the work of the design stage is performed well (c1), and the quality (v1) of the project tends to become high.

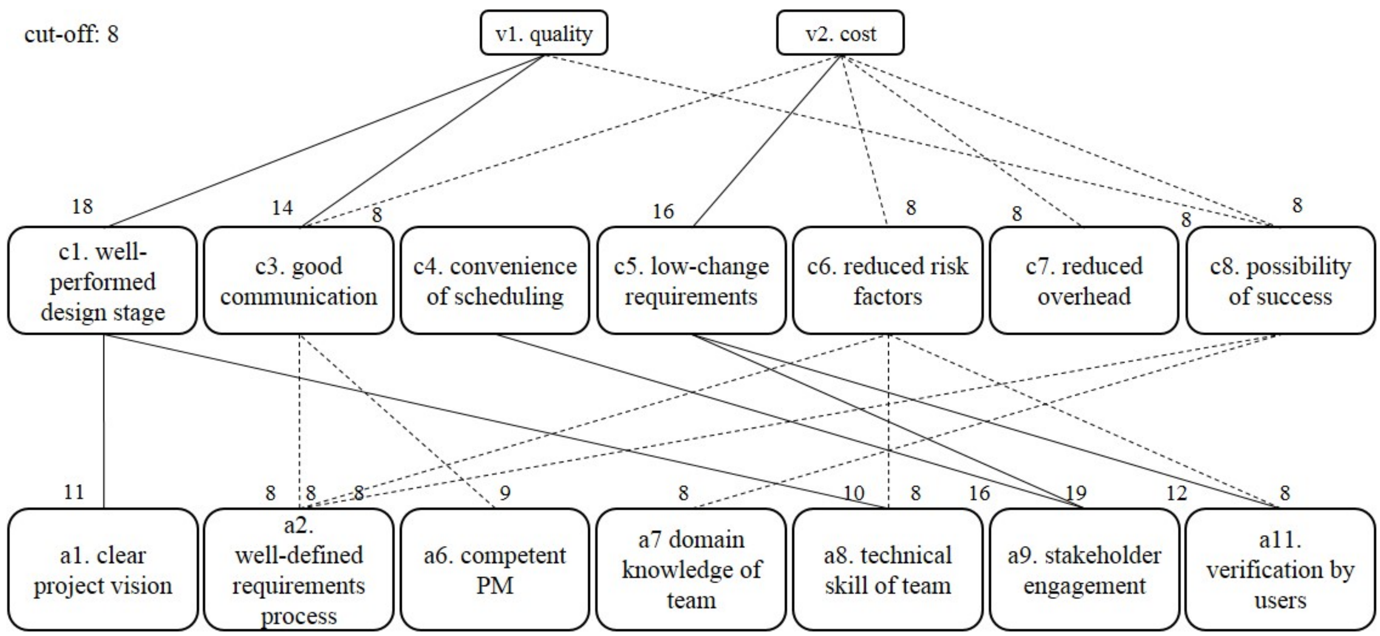

Fig. 2. HVM

B. Relation between projects and the attributes

We examine valid 99 projects; that is, we present the state of usage of the RET and investigate statistically whether there is a difference in the usage of RETs depending on the attributes of the project. First, Fig. 3 is a basic statistic that tells you how often each RET is used in Korea. In other words, it shows the percentages used in 99 projects for each RET, indicating that all the projects were interviewed regardless of the attributes of the project. Another overused RET is document analysis, which is performed in more than 90% of projects using document analysis. Contrary to these two RETs, observation and JAD are used only in less than 10% of projects. This subsection describes the statistical meaning of workshops with relatively similar use and nonuse ratios.

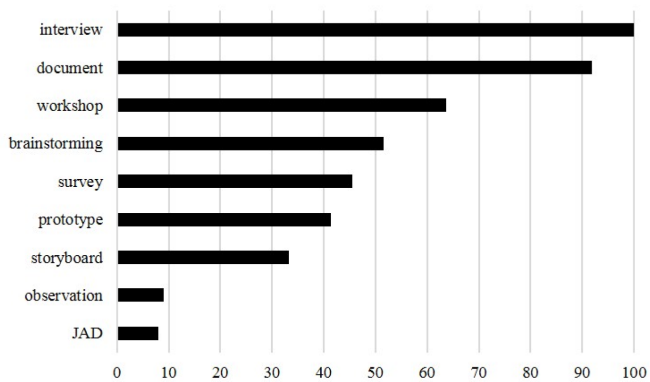

Fig. 3. Percentages of RET usage by project

Next, we analyze the cross-analysis using a chi-square test to determine whether each attribute has an association with the usage of RETs statistically.

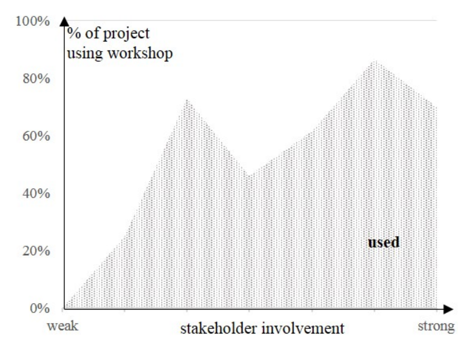

Fig. 4. Ratio of workshop usage depending on stakeholder engagement

Fig.4 shows the percentage of usage of workshops depending on the stakeholder engagement. While Fig.3 says that 63 of the 99 projects use workshops, if stakeholders participate inactively, they use the workshop technique less. We next test statistically whether the stakeholder engagement level and the usage of the workshop are independent. That is, we state the hypothesis H_0, *There is no association between the stakeholder engagement level and the usage of the workshop*. In Table I, the p-value is 0.039 (< 0.05), and so we reject the nulll hypothesis. Therefore, from the graph and χ^2 test results, we can expect Korean IT professionals to use workshops when stakeholder involvement is not low.

We also examine the results of binomial logistic regression analysis to test the strength of association between each attribute and the usage of the workshop. We choose a back-

TABLE I
χ^2 TEST

	value	df	asymptotic significance (2-way)
Pearson Chi-Square	13.292	6	0.039
Likelihood Ratio	13.91	6	0.031
Linear-by-Linear Association	3.662	1	0.056
N of Valid Cases	99		

ward elimination (condition) as the variable selection method because it is most commonly used. To find the best model, we perform a Hosmer-Lemeshow goodness of fit test. The smaller the χ^2 test is and the better the probability of the significance is, the better the model is. Therefore, we use the model achieved in step 9, for which the test yields a $\chi^2(9)$ of 2.081 which is insignificant (p=.978 $>$ 0.05), suggesting that the model is fit to the data well.

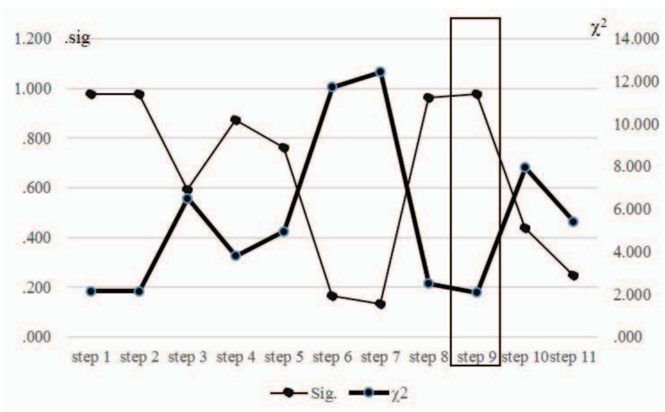

Fig. 5. The result of Hosmer-Lemeshow goodness of fit test

According to the model in step 9 (shown in Table II), the only attributes included in this logistic regression model are updated document (a3), amount of requirements (a5) and validation by users (a11). The more you participate in this requirement verification, the more projects you have participated in the workshop ($p < 0.05$). Other attributes are not related to the usage of workshops statistically.

TABLE II
THE RESULTS OF LOGISTIC REGRESSION

	B	S.E,	Wals	p
updated document	-0.32	0.166	3.697	0.055
amount of req	0.237	0.15	2.493	0.114
verification by user	0.569	0.16	12.626	0
constant	-1.512	0.849	3.177	0.075

Fig. 6 illustrates the decision tree, explaining how many projects having certain attributes used the workshop technique. Node 6 represents a case, in which the project has a large number of requirements and its user validates requirement. 92.5% projects on node 6 used workshop technique. In contrast to node 6, where workshops are most often used, the projects that likely use the technique less are determined by node 4. To arrive at node 4, users do not want to verify requirements at the first branch, and the vision of the project is clear from the start at the second branch.

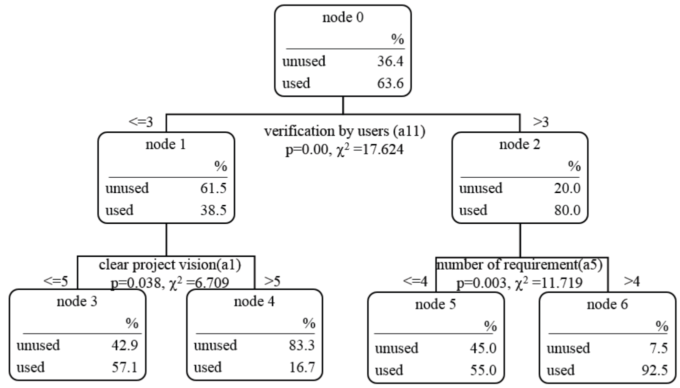

Fig. 6. Decision tree

Finally, we examine the prediction accuracy of the logistic regression model and of the decision tree model using the classification table shown in Table III. In the view of specificity, the logistic regression model predicts 18 (50%) of the 36 real workshop-based projects will use it while the decision tree model predicts 15 projects (42%). Therefore, the logistic regression model predicts a project that will not use the workshop better in this view. However, in the view of sensitivity, the logistic regression model and the decision tree model predict 54 (86%) and 60 (95%) of 65 projects using the workshop techqnique will use it, respectively. We usually want to determine whether a given project is likely to use a workshop technique, so the decision tree model, whose sensitivity is higher, is the better model. The classification accuracy of the logistic regression model and decision tree model is 73% and 76%, respectively.

TABLE III
CLASSIFICATION TABLE

logistic regression			
	predicted unused	predicted used	accuracy(%)
unused	18	18	50
used	9	54	86
total	37	63	73
decision tree			
	predicted unused	predicted used	accuracy(%)
unused	15	21	42
used	3	60	95
total	20	80	76

C. Example

This subsection illustrates how to apply our results to determine whether to use workshop techniques for a given new project. The elicitor evaluates each attribute of a new given project as a1 = 3, a2 = 7, a3 = 7, a4 = 5, a5 = 2, a6 = 7, a7 = 6, a8 = 3, a9 = 7, a10 = 2 and a11 = 3. If the values are applied to the logistic regression model and the probability of using a workshop is $\frac{e^{-1.512-0.32a_3+0.237a_5+0.569a_{11}}}{1+e^{-1.512-0.32a_3+0.237a_5+0.569a_{11}}}$, it is estimated to be $\frac{e^{-1.512-0.32\times7+0.237\times2+0.569\times7}}{1+e^{-1.512-0.32\times7+0.237\times2+0.569\times7}} = 0.669$, which is more than 0.5, and so it is highly likely to use a workshop. Next, we make a prediction using a decision tree. Because the value of a11 in node 0 is larger than 3, it branches to node 2, and the value of a5 is 2, which is smaller than 4, so a5 is selected for the current project. Therefore, 55% of the projects that have the same attributes have used the workshop technique, and it is recommended to use the technique.

In addition, you can use the HVM (Fig. 2) to predict the results and expectations of this project's attributes. In this project, attributes that are above average are a2, a3, a6, a7, a9 and a11. In this case, communication with stakeholders is smooth (c3), requirement changes are relatively small (c5), risk factors will also decrease (c8). Therefore, these attributes lead the consequence; that is more successfule projects (c11). That is, these attributes and consequences may cause the project to meet the value of cost rather than quality. For improving the quality, we need to identify the vision of the project before the design phase and reinforce the team with the required skills.

V. Conclusion

In this paper, we elicited the attributes influencing the decision of the RETs using laddering and analyzed the association between the usage and the attributes of real projects. In particular, we suggested how to build and interpret the logistic regression model and decision tree model using statistical results. This guideline shows the novice elicitor which RETs have been primarily used for the projects that have similar attributes to a given project.

Using the laddering technique, we could derive attributes that were suitable for the Korean IT culture and the operational definition of the attribute without limiting the existing research. This extracted attribute contains all the questions about elicitor, stakeholder, project, problem domain and the elicitation process proposed in previous research. In addition, we were able to build a ladder for the project's consequences and values based on the attributes of the project.

References

[1] Mike Mannion and Barry Keepence, "Smart requirements," *ACM SIGSOFT Software Engineering Notes*, vol. 20, no. 2, pp. 42–47, 1995.

[2] Khaled El Emam and A. Günes Koru, "A replicated survey of it software project failures," *IEEE software*, vol. 25, no. 5, 2008.

[3] Timo O.A. Lehtinen, Mika V. Mäntylän, Jari Vanhanen, Juha Itkonen, and Casper Lassenius, "Perceived causes of software project failures–an analysis of their relationships," *Information and Software Technology*, vol. 56, no. 6, pp. 623–643, 2014.

[4] Chun-Hsien Chen, Li Pheng Khoo, and Wei Yan, "A strategy for acquiring customer requirement patterns using laddering technique and art2 neural network," *Advanced Engineering Informatics*, vol. 16, no. 3, pp. 229–240, 2002.

[5] Neil AM Maiden and Gordon Rugg, "Acre: selecting methods for requirements acquisition," *Software Engineering Journal*, vol. 11, no. 3, pp. 183–192, 1996.

[6] Toshihiko Tsumaki and Tetsuo Tamai, "Framework for matching requirements elicitation techniques to project characteristics," *Software Process: Improvement and Practice*, vol. 11, no. 5, pp. 505–519, 2006.

[7] Saurabh Tiwari, Santosh Singh Rathore, and Atul Gupta, "Selecting requirement elicitation techniques for software projects," in *Proceedings of Software Engineering (CONSEG), 2012 CSI Sixth International Conference on*. IEEE, 2012, pp. 1–10.

[8] Dante Carrizo, Oscar Dieste, and Natalia Juristo, "Systematizing requirements elicitation technique selection," *Information and Software Technology*, vol. 56, no. 6, pp. 644–669, 2014.

[9] Maria-Isabel Sanchez-Segura, Fuensanta Medina-Dominguez, Diana-Marcela Vásquez-Bravo, Gustavo Illescas, and Cynthya García de Jesús, "Selecting a software elicitation technique according to layers of knowledge and preciseness: A case study," *Journal of Universal Computer Science*, vol. 23, no. 4, pp. 385–403, 2017.

[10] Nagy Ramadan Darwish, Ahmed Abdelaziz Mohamed, and Abdelghany Salah Abdelghany, "A hybrid machine learning model for selecting suitable requirements elicitation techniques," *International Journal of Computer Science and Information Security*, vol. 14, no. 6, pp. 380, 2016.

[11] Zheying Zhang, "Effective requirements development-a comparison of requirements elicitation techniques," in *Proceedings of 15th Software Quality Management Conference (SQM2007)*. British Computer Society, 2007, pp. 225–240.

[12] Masooma Yousuf and M Asger, "Comparison of various requirements elicitation techniques," *International Journal of Computer Applications*, vol. 116, no. 4, 2015.

[13] Tousif ur Rehman, Muhammad Naeem Ahmed Khan, and Naveed Riaz, "Analysis of requirement engineering processes, tools/techniques and methodologies," *International Journal of Information Technology and Computer Science (IJITCS)*, vol. 5, no. 3, pp. 40, 2013.

[14] Fares Anwar and Rozilawati Razali, "A practical guide to requirements elicitation techniques selection-an empirical study," *Middle-East Journal of Scientific Research*, vol. 11, no. 8, pp. 1059–1067, 2012.

[15] Asma Sajid, Ayesha Nayyar, and Athar Mohsin, "Modern trends towards requirement elicitation," in *Proceedings of the 2010 national software engineering conference*. ACM, 2010, p. 9.

[16] Reynolds, Thomas J, Olson, and Jerry C, *Understanding consumer decision making: The means-end approach to marketing and advertising strategy*, Psychology Press, 2001.

[17] J. Paul Peter and Jerry C. Olson, *Consumer Behavior*, McGraw-Hill Education, 3 edition, 2009.

[18] Reynolds, Thomas J, Gutman, and Jonathan, "Laddering theory, method, analysis, and interpretation," *Journal of advertising research*, vol. 28, no. 1, pp. 11–31, 1988.