

Evaluating a Lightweight Forum-based Tool: Empirical Studies on Requirements Elicitation Process

Han Lai^{1,2}

¹Chongqing Key Lab. of Electronic Commerce & Supply Chain System, Chongqing Technology and Business University, Chongqing, China

²School of Computer Science and Information Engineering, Chongqing Technology and Business University, Chongqing, China
Email: laihan_ctbu@126.com

Yuze Ni^{3,*}

³State Key Lab. of Software Engineering, Computer School, Wuhan University, Wuhan, China
Email: niyuze@whu.edu.cn

Abstract—The objective of this paper is to empirically evaluate SKLSEForum, a lightweight forum-based tool, for predicting the likelihood of acceptance of the tool in requirements elicitation practice and finding directions for improvement. We analyzed three data sets from a family of experiments, deriving from two replicated controlled experiments and a survey of expert panel. Results showed that SKLSEForum can improve the quantity of posts and judge satisfaction of requirements. In addition, the tool was perceived as useful and easy to use by the participants, who also expressed their intention to use SKLSEForum in the future. Furthermore, some insights about the improvement of SKLSEForum also have been found.

Index Terms—requirements engineering, requirements elicitation process, forum, expert panel, controlled experiment

I. INTRODUCTION

Open source and distributed software development communities already take advantage of Internet-based communication and coordination technologies to operate effectively. As social media has gained adoption, opportunities for creating software in new ways have risen to enhance and augment the old [1].

Ahmed, et al. [2] analyzed a dataset consisting of 1880 open source software projects covering a broad range of categories in this investigation. The results show that online forums play a significant role in managing software defects, implementation of new requirements and providing support to the users in open source software and have become a major source of assistance in maintenance of the open source projects.

However, many problems have emerged in the forum-based requirements elicitation [3], such as the disorder of post format, incompleteness of requirements information, low participation and enthusiasm of stakeholders, etc.

SKLSEForum [4] is a lightweight tool based on forum to support the forum-based requirements elicitation process. The tool includes five specialized features, i.e., thread template, reward topic, online chatting, thread state identification, and score exchange, with which it can complete and clarify the requirements' content, incentive the desire of participation.

This paper presents the empirical study, involving controlled experiment and survey of expert panel, defined and conducted to evaluate the SKLSEForum. We discuss the quantitative and qualitative findings of our empirical study and their potential for improving the SKLSEForum. These findings indicate that SKLSEForum could be considered as a promising tool for requirements elicitation.

This paper is organized as follows: Section 2 gives an overview of the five specialized features of our lightweight forum-based tool. Section 3 presents the evaluation method based on controlled experiments and a survey of expert panel, and how we conducted the empirical experiments. Section 4 gives the results of the experiments. In Section 5 we analyze the threats to the experiment validity. Section 6 presents the related work. Finally, Section 7 gives the conclusions and presents the future work.

II. SKLSEFORUM

P. Laurent and J. Cleland-Huang [3] explored and evaluated the forum-based requirements gathering and prioritization processes adopted by vendor-based open source software projects. They pointed that almost all of the forums they surveyed did a very poor job in managing the status of each feature request.

TABLE I
SKLSEFORUM'S FIVE FEATURES

Feature	Description	Requirements Activity	Functional Goal
Thread Templates	User can create a new opinion topic according to the guidance of requirements description template; and the topic must pass the template-based verification	Requirements Creation	Standardize the topic format, and improve the judge satisfaction of requirements
Reward-topic	The forum administrator can submit the reward topic to reward Q & A, and general users can answer the questions to prompt their reputation and increase scores	Requirements Improvement	Increase the reward topic's attention to gain more comments or suggestions, leading to higher judge satisfaction
Online Chatting	Users can chat with each other online in forum	Requirements Decision	Promote the communication and negotiation between users, leading to requirements with more judge satisfaction
Thread State Identification	The opinion topic has the marked status (e.g. New, Suggestion collected, Locked, and Unlocked, etc.). The transition flow among these states is based on authority and rules.	Requirements Management	Help the management of requirements, and instruct users to identify the status, leading to users' more participation on requirements decision to get higher requirements judge satisfaction
Score Exchange	General user can exchange the prize by their scores	The Whole Process	As an incentive mechanism to increase the users' participation and enthusiasm

Moreover, the stakeholders' participation is low because of the lack of two-way conversations which can engage more stakeholders in the requirements elicitation process. So it is necessary to provide a channel for stakeholders to communicate the process and decisions, seek clarification, etc.

To solve the problems P. Laurent and J. Cleland-Huang proposed, we established a framework ReqForum (see Fig. 1) to define the meta-model of the requirement elicitation forum, and developed a lightweight forum-based tool SKLSEForum with five specialized features (see Fig. 2) introduced in the requirements elicitation process [4]. The five features and their functional goals are introduced in Table I.

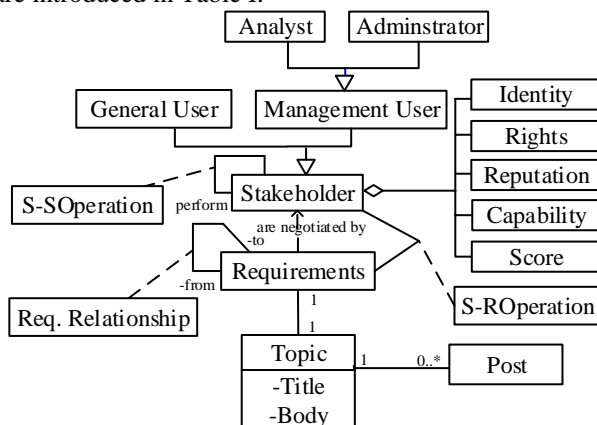


Figure 1. Forum-based requirement elicitation framework. [4]

In the forum-based requirement elicitation framework, stakeholders are classified into **General Users** and **Management Users**. The stakeholders who use the forum to give feedback on the open source communities' opinions or suggestions are deemed as **General Users**. Requirements analyst and the forum system administrator are deemed as **Management Users**. Each type of stakeholders has their own **Identity**, **Rights**, **Reputation**, **Capability** and **Score** in the forum. Different type of stakeholders has the different **Rights** to use forum's functions; the stakeholder's **Reputation** represents the trusted degree and also represents its prestige in the forum; the stakeholder's **Capability** represents the degree

of mastering domain knowledge, e.g. how to use the open source system's function or the applied domain of open source system used; the stakeholder's **Score** represents its active degree which participate in the forum's discussion. The **Requirements** are represented as threads. A requirement is corresponding to a **Topic** and multiple **Posts**. A general **Topic** consists of title and body. The **Req. Relation** represents the relationship, which exists between the requirements. The **S-ROperation** represents the stakeholders can perform various operations on the forum's topic and post to reflect the activity of different stakeholders to get involved in requirement elicitation. The **S-SOperation** represents the operations between Stakeholders, e.g. management user rewards score to general users in the forum to discuss, stakeholders communicate with each other by sending and receiving message etc.

Different from the ordinary forum, a requirement elicitation process can be supported by SKLSEForum. The process is shown in Fig. 3. Firstly, the stakeholders can create requirements by using the thread templates; secondly, in the phrase of requirements improvement, the stakeholders comment on the posts, and the Analysts can set reward topic to draw more general user's attention and participate the feedback. Then, during the requirements decision, the online chatting provides a platform for stakeholders to communicate with each other in the synchronous setting. In addition, the thread state identification can help manage the requirements and prompt the awareness about the requirements' disposition. The score exchange can stimulate stakeholders to participate the discussion.

By using the five specific features in SKLSEForum, we can join these features into the SECI model [5]. The thread templates transform the stakeholders' ideas into formative requirements and make other stakeholders understand more about the requirements, thus it facilitates the stakeholders' knowledge Externalization and

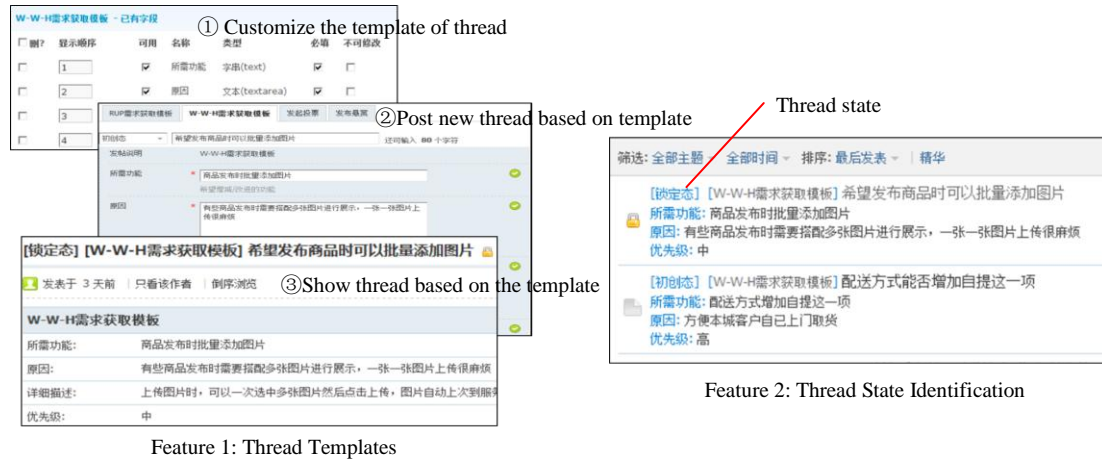
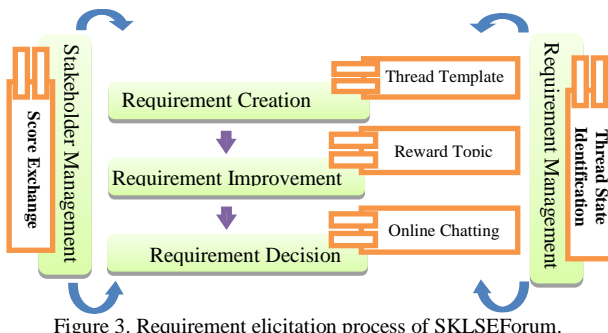


Figure 2. Screenshot of the main features of SKLSEForum in Chinese.



Internalization. Reward topic combines the comments from different stakeholders, and facilitates the Combination. Thread state identification can instruct users to identify the statuses of requirements and score exchange can stimulate more stakeholders participate, thus it can facilitate the Socialization.

III. SKLSEFORUM EVALUATION

This section describes the empirical validation of SKLSEForum. This was done by conducting two replicated controlled experiments and a survey of expert panel.

A. Evaluation Design

The evaluation design refers to the guidelines proposed by Wohlin [6]. According to the Goal-Question-Metric (GQM) template [7], the goal of the experiment is to test

the SKLSEForum with respect to its performance and perception for predicting the likelihood of acceptance of the tool in requirements elicitation practice and find directions for improvement.

The research questions addressed by the experimentation are:

RQ1: Can the performance of requirements elicitation be improved by applying SKLSEForum?

RQ2: Is SKLSEForum perceived as both easy to use and useful?

RQ3: Is there an intention to use SKLSEForum in the future?

RQ4: What improvement can be made for SKLSEForum?

The first research question (RQ1) was addressed by defining the following hypotheses:

H1₀: SKLSEForum will not improve the quantity of posts.

H2₀: SKLSEForum will not improve the judge satisfaction of requirements.

The second research question (RQ2) was addressed by the formulation of the following hypotheses:

H3₀: SKLSEForum is perceived as not useful.

H4₀: SKLSEForum is perceived as difficult to use.

The third research question (RQ3) was addressed by defining the following hypotheses:

H5₀: There is no intention to use SKLSEForum in the future.

Table II lists all the dependent variables in the experiments. The chosen Perceived-based dependent variables are based on the adapted Davis's TAM [8][9]. The TAM is one of the most widely applied theoretical models to study user acceptance and usage behavior of emerging information technologies, and it has received extensive empirical support through validations and replications [10][11].

TABLE II
DEPENDENT VARIABLES

Variable name	Measure
Quantity of Posts	Count
Judge Satisfaction	Calculated as the mean of the grades from two teachers
Perceived Usefulness	Calculated as the mean of the items obtained from the questionnaire
Perceived Ease of Use	Calculated as the mean of the items obtained from the questionnaire
Self-predicted Future Use	Calculated as the mean of the items obtained from the questionnaire

B. Questionnaire Design

To obtain the opinions of experts and students in the experiment groups about the perceived usefulness, perceived ease of use and self-predicted future use of the whole SKLSEForum, we designed the questions in the questionnaire based on the adapted Davis's TAM [12], which has six items in perceived usefulness, six items in perceived ease of use and two items in self-predicted future use, as shown in Table III. The questionnaire also included five semi-open questions to inquire the subjects whether the five specific features could achieve their functional goals in Table I, and one open question to collect ideas, advice or suggestions to SKLSEForum for improvement in the future.

C. Participants and Training

28 fourth-year Computer Science undergraduate students at the Chongqing Technology and Business University participated in the controlled experiments. The participants are volunteers and we provide a movie ticket as the reward of participation. In order to avoid persistence effects, the students are chosen from those who had never done similar experiments before. The students were randomly divided into 4 groups, (i.e. two experiment groups, and two control groups), each composed of 7 members.

The subjects respectively received 80 minutes training on SKLSEForum or ordinary forum (i.e. Discuz¹, the most widely used forum application in China and is the SKLSEForum's prototype). The training was consisted of 40 minutes presentation for explaining the main characteristics of SKLSEForum or Discuz, and 40 minutes for tool demo. They did not receive further teachers' assistance while they were using the forum.

We invited 16 experts from different professional backgrounds to participate our survey by email and 14

TABLE III
THE QUESTIONNAIRE DESIGN

Code	Measurement standard
Perceived Usefulness	
U1	Using this collaborative tool in my job, I would be able to accomplish requirements elicitation tasks more quickly.
U2	Using this tool would improve my performance on requirements elicitation tasks.
U3	Using this tool for requirements elicitation tasks would increase my productivity.
U4	Using this tool would enhance my effectiveness on requirements elicitation tasks.
U5	Using this tool would make it easier to do requirements elicitation tasks.
U6	I would find this tool useful to perform requirements elicitation tasks.
Perceived Ease of Use	
E1	Learning to operate this tool would be easy for me.
E2	I would find it easy to get this tool to do what I want it to do to perform most of the requirements elicitation tasks.
E3	My interaction with this tool would be clear and understandable.
E4	It would be easy to become skillful in using this tool.
E5	It would be easy to remember how to perform various requirements elicitation tasks using this tool.
E6	I would find this tool easy to use.
Self-predicted Future Use	
S1	Assuming this tool would be available on my job, I predict that I will use it on a regular basis in the future.
S2	I would prefer using this tool for requirements elicitation tasks.

accepted (representing a take-up rate of 88%), including requirements engineering (RE) researchers, experienced open source community (OSC) forum managers, active forum users and open source software consultants. In addition, we also included experts from collaborative computing. Table IV shows the profile of the experts who participated in the study.

TABLE IV
DEMOGRAPHICS OF 14 EXPERTS

	Roles	Years of Experience	Research Area	Position / Relevant Experience
1	A	6	RE	Associate Professor
2	A	8	OSC and RE	Professor
3	A	4	RE and CC	Associate Professor
4	A	3	RE	Associate Professor
5	A	10	OSC and RE	Professor
6	P	5	OSC	Forum Manager
7	P	4	OSC	Forum Moderator
8	P	7	OSC and RE	IT Consultant
9	P	9	OSC	IT Consultant
10	P	6	OSC	IT Consultant
11	P	5	OSC	Business Analyst
12	B	3	OSC and RE	Forum Moderator
13	B	4	OSC and RE	Forum Manager
14	B	3	OSC and CC	Active User
A: Academic P: Practitioner B: Both CC: Collaborative Computing OSC: Open Source Community RE: Requirement engineering				

The experts are given some materials prepared with the questionnaire. The prepared materials were an instructional video describing the main characteristics of SKLSEForum, and the procedure for applying it.

¹ Discuz: <http://www.discuz.net>

D. Controlled Experiment & Survey

Following the general plan of evaluation, we carried out two controlled experiments and a survey, as is shown in Fig. 4.

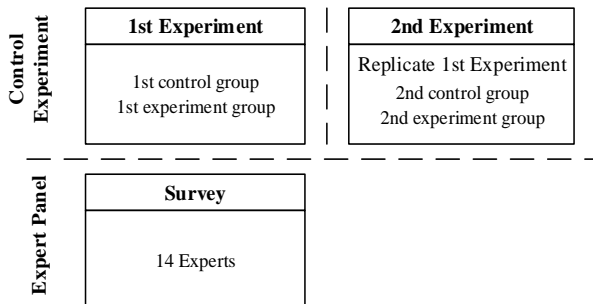


Figure 4. Overview of the evaluation.

The evaluation process involves controlled experiment and survey, depending on the type of participants and the experimental design employed.

Controlled experiment: Undergraduate students were divided into four groups of 7 members. Each experiment consisted of two groups, and the second experiment is a strict replication of the first.

Survey: The survey consists of the 14 experts.

In order to discover potential problems in the questionnaire design and application, we ran a pilot study involving three researchers in the fields of open source community and RE. Through analyzing the pilot test responses, we made changes as a result of the feedback. We also used the pilot study to assess the time required to complete the questionnaire which we estimated to be 10 min.

In Controlled experiment, we conducted two replicated controlled experiments under the same conditions (strict replication), changing only the students [13]. The control groups use Discuz as ordinary forum, while experiment group use SKLSEForum. Each of the teams was asked to design and document software architecture of a “university library web portal” for three days. The second experiment was performed on the same three day as the first one. This application scenario was deemed sufficiently complex such that students would be able to generate requirements for three days and also because it is highly familiar to college students. The teams were asked to use the tools in the process of requirements elicitation, to capture the key business scene, case, questions and so on. During the experiment, the method of role play [14] was adopted and the students used the forum by means of remote access.

After three days, we collected the requirements the students proposed, and the quantity of posts each proposed in four groups is sorted and counted; the judge satisfaction value of requirements is graded by teachers. Moreover, the two experiment groups were asked to finish the questionnaires after the experiment. In order to prevent the potential bias, the students were told that their answers would be treated anonymously, and they need

not please the experimenters by favorable judgments of SKLSEForum. Moreover, before the experiment, the students were also informed that their performance in the experiment would not affect them or their grade.

In the survey of expert panel, we handed out the questionnaires by E-mail, and collected the feedback from the experts about the perceived usefulness, perceived ease of use and self-predicted future use of SKLSEForum.

IV. RESULTS

This section reports the analysis result of the data collected during the studies. First, we assess whether the SKLSEForum can improve the quantity of posts and judge satisfaction of the requirements elicited. Then, we investigate perception-based variables (i.e. perceived useful, ease of use, self-predicted future use) the results from the students’ operation experience and experts’ knowledge.

The quantitative analysis was performed by using the SPSS v19 statistical tool and $\alpha=0.05$.

A. Performance-based Measures

After three days of controlled experiments, we collect the whole requirements elicited from the four group students. Table V shows the mean quantity of posts by each student contributed in the four experiments.

TABLE V
MEAN QUANTITY OF POSTS

Experi ment	Group	Mean	SD	Statistical analysis results
1st Experi ment	1st Control Group	8.29	2.498	$p=0.010$ (t-test)
	1st Experiment Group	12.71	3.592	
2nd Experi ment	2nd Control Group	8.71	1.704	$p=0.005$ (t-test)
	2nd Experiment Group	13.57	3.867	

The data from Table V highlight that the students in experiment group has generated more posts that represent the key business scenarios, use cases, and problems than control group. The data are normally distributed, therefore, one-tailed two independent sample t-test are adopted.

The comparison of the mean quantity of posts highlights that the SKLSEForum can easily promote the effectiveness of the stakeholders’ participation and stimulates their enthusiasm. By analyzing the posts elicited from the experiment groups, it seems that stakeholders is encouraged to generate more requirements in order to get redeem gifts, and based on the online chatting, more opinions and ideas are captured.

The two teachers performed a general analysis of the requirements, to determine the judge satisfaction value. Table VI represents the judge satisfaction value of the requirements elicited from the four group students.

TABLE VI
JUDGE SATISFACTION VALUE

Experiment	Group	T1	T2	Judge Satisfaction Value
1st Experiment	1st Control Group	62	70	66
	1st Experiment Group	81	87	84
2nd Experiment	2nd Control Group	59	67	63
	2nd Experiment Group	76	85	80.5
T1: 1st Teacher T2: 2nd Teacher				

It is obvious that the judge satisfaction of requirements in experiment groups (i.e. the 1st experiment group and 2nd experiment group) got higher values than the control groups (i.e. the 1st control group and 2nd control group). In conclusion, the analysis of the data collected during the experiment indicate that using SKLSEForum can improve the requirements' judge satisfaction during the requirements elicitation process in a global environment. By analyzing the requirements elicited from the experiment groups, it seems that stakeholders are more likely to submit requirements with more judge satisfaction by using SKLSEForum.

B. Perception-based Measures

The construct validity of the metrics with the perceived usefulness, perceived ease of use and self-predicted future use was evaluated by using inter-item correlation analysis. The evaluation was based on two criteria: convergent and discriminant validity. Low divergent validity values indicate high discriminant validity. According to Campbell and Fiske [15] an item's convergent validity value must be higher than its divergent validity value; otherwise the data on the item should not be used. The results of inter-item correlation analysis (see Table VII for an example) for the three data sets were:

- 1st Experiment Group: E3 and E6 did not pass the validity test. E3 and E6 were therefore excluded from the analysis. With the exclusion of this item, the average convergent validity was 0.60 for perceived usefulness, 0.43 for perceived ease of use, and 0.77 for self-predicted future use.
- 2nd Experiment Group: S2 did not pass the validity test. With the exclusion of this item, the average convergent validity was 0.49 for perceived usefulness, 0.40 for perceived ease of use, and 0.53 for self-predicted future use.
- Expert panel: E2 and U3 did not pass the validity test. With the exclusion of this item, the average convergent validity was 0.52 for perceived usefulness, 0.45 for perceived ease of use, and 0.64 for self-predicted future use.

The use of multiple items to measure a same construct requires the examination of the reliability or internal consistency of the questionnaire. We calculated the Cronbach's alpha for each set of questions in the questionnaire. For this analysis, the items that did not pass the validity test were excluded from their data sets. Table VIII indicates that almost all the questions have an alpha value greater than 0.7, which is a common reliability threshold [16], the least reliable one is 'Perceived Ease of Use' of 1st Experiment Group.

Table IX shows descriptive statistics about perceived usefulness, perceived ease of use and self-predicted future use of SKLSEForm provided by the experiment group students and expert panel. The result shows that we corroborated empirically the students and experts perceived SKLSEForm as being useful and easy to use, and that there is an intention to use it in the future.

In order to analyze the actual usefulness and ease of use of five specific features in SKLSEForum, the students were asked to write down their opinions based on the actual using experience in SKLSEForum.

The students' response showed some disagreement that score exchange feature is useful. One reason could be that potential stakeholders could not realize the function of the scores, and are regardless of the exchange gift or virtual rewards. Another one could be that since the period of experiment is short, thus the stakeholders cannot accumulate many scores in just three days to exchange gift or virtual rewards. The few opportunity for stakeholders to use the score exchange feature lead to the students' disagreement with the usefulness of the feature.

Some students indicated that the thread templates feature and thread state identification feature are not easy to use as expected. For the thread templates feature, a factor that might reduce the ease of use is that the usability of the thread templates is restricted to the experienced users only. Templates can conduct the mandatory and optional items for the requirements information, but for the novice, the feature is not easy to apply especially when they set up simple requirements. And a mass of templates will cause a waste of time and aggrandize redundancy of requirements information, making the thread templates feature not easy to use. On the other hand, the thread state identification feature is mainly applied by forum administrators or project managers to add or alter the requirements states. Similarly for the novice, it is not easy to distinguish or identify the requirement states, let alone conducting the requirement elicitation process according to the requirements states.

In the structured questionnaire conducted during the expert panel, the experts were asked to feedback their views and suggestions about the SKLSEForum.

Some experts pointed out that by applying the SKLSEForum, it is unavoidable to increase the dependence degree on stakeholders' participation. Despite this conforms to our intention of promoting the effectiveness of the stakeholders' participation, one expert said "This will increase the management cost of the open forum". One practitioner also remarked that "In open forum, the requirements provided by users are imperfect and still need project manager to collect and consolidate". Another manager said that "The topic status is not necessary for users, but is more important for site manager".

Some experts went further and suggested other features to incentive the stakeholders' participation: "You can show the users ranking based on the contribution

TABLE VII
INTER-ITEM CORRELATION ANALYSIS FOR THE FIRST EXPERIMENT GROUP

		Perceived Usefulness						Perceived Ease of Use						Self-predicted Future Use		Overall		
		U1	U2	U3	U4	U5	U6	E1	E2	E3	E4	E5	E6	S1	S2	CV	DV	Valid?
Perceived Usefulness	U1	1.00														0.65	0.39	Y E S
	U2	0.44	1.00													0.47	0.46	Y E S
	U3	0.65	0.81	1.00												0.71	0.38	Y E S
	U4	0.65	0.75	0.95	1.00											0.70	0.36	Y E S
	U5	0.68	0.00	0.44	0.51	1.00										0.45	0.32	Y E S
	U6	0.84	0.37	0.72	0.63	0.61	1.00									0.63	0.41	Y E S
Perceived Ease of Use	E1	0.55	-0.40	0.35	0.23	0.31	0.76	1.00								0.43	0.38	Y E S
	E2	0.68	0.60	0.59	0.51	0.00	0.61	0.62	1.00							0.47	0.40	Y E S
	E3	0.42	0.00	0.54	0.63	0.82	0.50	0.38	0.00	1.00						0.27	0.40	N O
	E4	0.40	-0.44	-0.04	0.05	0.34	0.42	0.73	0.34	0.42	1.00					0.35	0.27	Y E S
	E5	0.26	0.71	0.36	0.26	-0.44	0.27	0.06	0.66	-0.54	-0.26	1.00				0.46	0.42	Y E S
	E6	0.59	0.78	0.76	0.59	0.00	0.71	0.54	0.87	0.00	0.00	0.76	1.00			0.43	0.48	N O
Self-predicted Future Use	S1	0.11	-0.12	-0.09	-0.27	0.19	0.00	0.20	0.00	0.00	-0.11	0.31	0.00	1.00		0.77	0.12	Y E S

TABLE VIII
QUESTIONNAIRE CRONBACH'S ALPHA

Group	Item	Cronbach's α
1st Experiment Group	Perceived Usefulness	0.882
	Perceived Ease of Use	0.694
	Self-predicted Future Use	0.851
2nd Experiment Group	Perceived Usefulness	0.892
	Perceived Ease of Use	0.755
	Self-predicted Future Use	0.789
Expert panel	Perceived Usefulness	0.725
	Perceived Ease of Use	0.812
	Self-predicted Future Use	0.732

each user made, by computing their requirements' quantity of posts and judge satisfaction." Moreover, one researcher suggested replacing the score exchange feature by "using people's competitive psychology in one group or organization, like user level classification".

V. THREATS TO VALIDITY

In the following, we explain how we dealt with the validity threats.

Construct validity: By using the convergent and discriminant validity, we eliminated the invalid data in the feedback of questionnaire. The internal consistency of the questions was verified by using the Cronbach's α which is over 0.7 for all items instead for one, which is higher >0.6 and that is valid for exploratory studies [17].

Internal validity: The experts had enough maturity level to participate in the expert panel due to their previous knowledge and experience of Open Source Community, Requirements Engineering and Cooperative Computing, thus the responses provided by the experts on the SKLSEForum can be biased by the personal experience and domain knowledge. The teachers who evaluate the requirements quantity of posts and judge satisfaction face the same question, so we employed two teachers who come from third-party to avoid the bias. Moreover, the results may be influenced by the students' ability and energy kept during the experiment period. We have reduced the students' bias by informing the students

that their answers would be treated anonymously, and their performance in the experiment would not affect them or their grade.

TABLE IX
DESCRIPTIVE STATISTICS FOR QUESTIONNAIRES

Data set	Item	Min	Max	Mean	SD	Statistical analysis results
1st Experiment Group	Perceived Usefulness	2.67	4.67	3.74	0.751	$P=0.020$ (t-test)
	Perceived Ease of Use	3.17	4.50	4.00	0.451	$p=0.000$ (t-test)
	Self-predicted Future Use	3.00	5.00	4.14	0.748	$p=0.004$ (t-test)
2nd Experiment Group	Perceived Usefulness	3.50	4.33	4.01	0.343	$p=0.013$ (t-test)
	Perceived Ease of Use	3.33	4.33	4.02	0.390	$p=0.002$ (t-test)
	Self-predicted Future Use	3.50	4.50	4.00	0.500	$p=0.011$ (t-test)
Expert panel	Perceived Usefulness	3.21	4.59	4.29	0.501	$p=0.001$ (t-test)
	Perceived Ease of Use	3.14	4.48	4.17	0.467	$p=0.010$ (t-test)
	Self-predicted Future Use	3.42	4.71	4.30	0.658	$p=0.006$ (t-test)

External validity: Since we have evaluated the applicability of the SKLSEForum with different kind of stakeholders, i.e. students, researchers and practitioners, our results can have general validity. According to Tichy [18], the students are appropriately trained and the data is used to establish a trend, thus here we consider the students' ability for requirements elicitation is comparable to that of typical novice requirements analyst [19]. However, we used the "university library web portal" to conduct the experiment and collect the data. To show that our findings are true for other cases we should

test the applicability of the SKLSEForum to other people in other industrial contexts.

Conclusion validity: An important threat to the conclusion validity of our study is the relatively small number sample that participated to the studies. In respect to the random heterogeneity of participants, the participants have received the same training about SKLSEForum. We should organize other study to have a bigger data sample to draw our conclusions.

VI. RELATED WORK

There are several researches related our work. One of them is studies on forum based requirements elicitation. Cleland-Huang et al. explore and evaluate the forum-based requirements gathering and prioritization processes by vendor-based open source software projects [3]. The study reports a number of interesting lessons that can be learned and applied in forum-based requirements elicitation processes. Then they proposed a method to improve recommendations by enhancing stakeholder profiles [20], utilized recommender systems to support software requirements elicitation [21], applying a recommender system for requirements elicitation in large-scale software projects [22], and provided automated support for managing feature requests in open forums [23]. They highlighted the methods like data mining and machine learning techniques which can give aid to the software requirements elicitation. However they didn't provide an integration tool to solve the problems occurred in open forums. In addition, none of the mentioned studies has been evaluated in practical industry context.

In the field of requirements engineering case tools, many requirements elicitation and management tools have been proposed from academic, like specialized KAOS' support tool Objectiver [24], Requirements engineering specific wikis [25]. One drawback of these tools is that they general lack proper empirical evaluations and very few have had widespread use in practice. Even in the implement of empirical evaluation, they have given a comprehensive description of their usage context, study design, and discussed their validity, but their usage contexts are most in academia. If further evaluation can be carried out in industry, it will increase the possibility of transferring it to industry context [26][27]. Here we adopted on the expert panel to capture their perceived usefulness, perceived ease of use and self-predicted future use in industry context.

VII. CONCLUSION AND FUTURE WORK

The main goal of this study is to evaluate SKLSEForum for supporting the requirements elicitation process and understand the demands for adjusting and extending that tool.

Through the two replicated controlled experiments, we have evaluated that the application of SKLSEForum can improve both the quantity of posts and judge satisfaction of the requirements elicited. Furthermore, by analyzing the responses from the expert panel and the students in

experiment groups, we evaluated the perceived usefulness, perceived ease of use and self-predicted future use of SKLSEForum. The results have found SKLSEForum is being useful and easy to use, and that there is an intention to use it in the future.

In addition, the qualitative analysis indicated that the feature with more potential is the score exchange. On the other hand, the students indicated that users may have some difficulties using SKLSEForum, because SKLSEForum is a prototype tool that still needs improvement in its usability. The thread templates and thread state identification are considered as hard features to use. Moreover, the questionnaire results and attitudes of the experts towards SKLSEForum are likely to be of interest to a wide range of specialists in both the research community and the software industry.

The evaluation show us the improvements can be made to SKLSEForum, specifically the ones related to improve the ease of use of SKLSEForum. In our future work, we plan to improve our SKLSEForum, and conduct experiments with practitioners in industry contexts to acquire more empirical results. Other research directions we will investigate how to increase the perceived usefulness, perceived ease of use and self-predicted future use about integrate our tool as a plug-in into industry-level requirements engineering case tool and the usability in real world. For the thread templates provided by SKLSEForum, we plan to find a method to improve the templates structure for easier knowledge externalization, to inspire the expression of stakeholders, like the templates based on story [28]. Furthermore, we will introduce the argumentative discussions based on IBIS [29] and SIOC [30] to promote the stakeholders' discussions in the SKLSEForum.

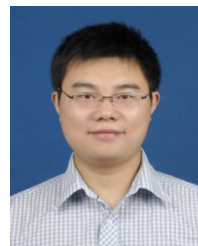
ACKNOWLEDGMENT

The research was supported by the National Natural Science Foundation of China under Grant No. 61170026, the Chongqing Key Lab. of Electronic Commerce & Supply Chain System Special Fund (2012ECSC0210).

REFERENCES

- [1] A. Begel, R. DeLine and T. Zimmerman, "Social Media for Software Engineering," in *Proceedings of the FSE/SDP workshop on Future of software engineering research*, 2010, pp. 33-38.
- [2] F. Ahmed, P. Campbell, A. Jaffar, and L. F. Capretz, "Myths and Realities about Online Forums in Open Source Software Development: An Empirical Study," *Open Software Engineering Journal*, vol. 4, pp. 52-63, 2010.
- [3] P. Laurent and J. Cleland-Huang, "Lessons Learned from Open Source Projects for Facilitating Online Requirements Processes," in *Proceedings of the 15th International Working Conference on Requirements Engineering: Foundation for Software Quality (REFSQ'09)*. Berlin, Heidelberg: SpringerVerlag, 2009, pp. 240-255.
- [4] H. Lai, R. Peng, D. Sun, and J. Liu, "A lightweight forum-based distributed requirement elicitation process for open source community," *International Journal of Advancements in Computing Technology*, vol. 4, pp. 138-145, 2012.

- [5] Nonaka I, Takeuchi H, Takeuchi H. *The knowledge-creating company: how Japanese companies create the dynamics of innovation*. Oxford University Press, USA, 1995.
- [6] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in software engineering*: Springer, 2012.
- [7] V. R. Basili, and H. D. Rombach, "The TAME project: Towards improvement-oriented software environments," *IEEE Transactions on Software Engineering*, vol. 14, no. 6, pp. 758-773, 1988.
- [8] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS quarterly*, pp. 319-340, 1989.
- [9] Su-pi Su, Chung-hung Tsai and Wei-lin Hsu. "Extending the TAM Model to Explore the Factors Affecting Intention to Use Telecare Systems," *Journal of Computers*, vol. 8, no. 2, pp. 525-532, 2013.
- [10] V. Venkatesh, "Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model," *Information systems research*, vol. 11, pp. 342-365, 2000.
- [11] S. Quan, C. Hao, and Y. Jianxin, "Factors Influencing the Adoption of Mobile Service in China: An Integration of TAM," *Journal of computers*, vol. 5, no. 5, pp. 799-806, 2010.
- [12] Babar, Muhammad Ali, Dietmar Winkler, and Stefan Biffl. "Evaluating the usefulness and ease of use of a groupware tool for the software architecture evaluation process." In *Proceedings of The First International Symposium on Empirical Software Engineering and Measurement*, 2007.
- [13] V. R. Basili, F. Shull and F. Lanubile, "Building knowledge through families of experiments," *IEEE Trans. on software Engineering*, vol. 25, pp. 456-473, 1999.
- [14] D. Zowghi and S. Paryani, "Teaching requirements engineering through role playing: Lessons learnt," in *11th IEEE International Requirements Engineering Conference*, 2003.
- [15] D. T. Campbell, and D. W. Fiske, "Convergent and discriminant validation by the multitrait-multimethod matrix," *Psychological bulletin*, vol. 56, no. 2, pp. 81-105, 1959.
- [16] J. Nunally, *Psychometric Theory*, second ed., McGraw-Hill, New York, NY, 1978.
- [17] D. George and P. Mallery. *SPSS for Windows Step-by-Step*. Allyn & Bacon, Inc., 8th edition, 2007.
- [18] W. F. Tichy, "Hints for reviewing empirical work in software engineering," *Empirical Software Engineering*, vol. 5, no. 4, pp. 309-312, 2000.
- [19] M. Höst, B. Regnell, and C. Wohlin, "Using students as subjects—a comparative study of students and professionals in lead-time impact assessment," *Empirical Software Engineering*, vol. 5, no. 3, pp. 201-214, 2000.
- [20] C. Castro-Herrera, J. Cleland-Huang, and B. Mobasher, "Enhancing Stakeholder Profiles to Improve Recommendations in Online Requirements Elicitation", In *Proceedings of the 17th IEEE International Conference on Requirements Engineering*, pp. 37-46, 2009.
- [21] C. Castro-Herrera and J. Cleland-Huang, "Utilizing recommender systems to support software requirements elicitation", In *Proceedings of the 2nd International Workshop on Recommendation Systems for Software Engineering*, pp. 6-10, 2010.
- [22] C. Castro-Herrera, C. Duan, J. Cleland-Huang, and B. Mobasher, "A recommender system for requirements elicitation in large-scale software projects", In *Proceedings of the ACM symposium on Applied Computing*, pp. 1419-1426, 2009.
- [23] J. Cleland-Huang, H. Dumitru, C. Duan, and C. Castro-Herrera, "Automated support for managing feature requests in open forums", *Communications of the ACM*, vol. 52, pp. 68-74, 2009.
- [24] H. S. Al-Subaie and T. S. Maibaum, "Evaluating the effectiveness of a goal-oriented requirements engineering method," in *Fourth International Workshop on Comparative Evaluation in Requirements Engineering*, 2006, pp. 8-19.
- [25] R. Peng and H. Lai, "DRE-specific Wikis for Distributed Requirements Engineering: A Review," in *19th Asia-Pacific Software Engineering Conference*, 2012, pp. 116-126.
- [26] M. Ivarsson and T. Gorschek, "Technology transfer decision support in requirements engineering research: a systematic review of REj," *Requirements engineering*, vol. 14, pp. 155-175, 2009.
- [27] F. Ahmed, "Software Requirements Engineer: An Empirical Study about Non-Technical Skills," *Journal of Software*, vol. 7, no. 2, pp. 389-397, 2012.
- [28] V. Laporti, M. R. Borges, and V. Braganholo, "Athena: A collaborative approach to requirements elicitation," *Computers in Industry*, vol. 60, no. 6, pp. 367-380, 2009.
- [29] W. Kunz, and H. W. Rittel, *Issues as elements of information systems*: Institute of Urban and Regional Development, University of California Berkeley, California, 1970.
- [30] C. Lange, U. Bojars, T. Groza, J. Breslin, and S. Handschuh, "Expressing argumentative discussions in social media sites," in *Workshop on Social Data on the Web*, 2008.



software process improvement and empirical software engineering.



Yuze Ni was born in 1991. She received bachelor degree in Wuhan University in 2012. She is now a master student in Wuhan University. Her research interests include requirement engineering and empirical software engineering.