Question's Advisor
A Wizard Interface to Teach Novice Programmers How to Post "Better" Questions in Stack Overflow

José Remígio[1], Franck Aragão[1], Cleyton Souza[1,2], Evandro Costa[3] and Joseana Fechine[2]

[1]Federal Institute of Education, Science and Technology of Paraíba, IFPB, Monteiro, Paraíba, Brazil
[2]Federal University of Campina Grande, UFCG, Campina Grande, Paraíba, Brazil
[3]Laboratory of Artificial Intelligence, LIA, Campina Grande, Paraíba, Brazil

Keywords: Stack Overflow, Answerability, System, Question Quality.

Abstract: Programmers often recur for online communities in order to find help for a current problem that they are facing. However, after sharing a question, its author has no guarantee if he will receive an answer, neither when. Recent studies have found that low quality is one of the top reasons why questions remain unanswered. In this work, we conducted a qualitative study aiming identifying what programmers are looking in a question that they decide to answer. Based on this feedback, we designed a tool to help programmers to write high quality questions. We named the app Questions' Advisor, due his role of helping but without forcing the user to follow it, and it is available for desktop and mobile clients. We believe it could be very helpful, especially for novice programmers.

1   INTRODUCTION

Stack Overflow is a Community Question and Answering (CQA) site for professional and enthusiast programmers. It is built as part of the Stack Exchange platform of Q&A sites and it is the largest community of the network with 6 million users and over 12 million asked questions[1]. Stack Overflow works like a thread based community. The questions' titles are presented in a "wall", ordered by their latest interaction, and, by clicking them, users can see the description and answer it.

In general, questions in Stack Overflow are answered in a very short time (Mamykina Manoim and Mittal, 2011). However, after sharing a question, its author has no guarantee if he will receive an answer, neither when. According Hao, Shu and Irawan (2014), over the years, the number of unanswered or ignored questions is constantly increasing. Interestingly, the fact that those questions are not answered is not caused by users not having seen them (Baltadzhieva and Chrupała, 2015). One of the main reasons to a question being

ignored is its low quality (Asaduzzaman *et al.*, 2013).

Some recent studies have found that there is a correlation between the question characteristics and its responsiveness. In Facebook, for example, Teevan, Morris and Panovich (2011) found that a concise style of question-asking, a defined scope (or audience), and the inclusion of a question mark were associated with more and higher quality responses within shorter periods of time. Regarding Stack Overflow, many works suggest that the quality of the question itself can have an important effect on the likelihood of getting useful answers (Baltadzhieva and Chrupała, 2015).

In this work, we asked programmers which characteristics they were expecting when they choose a question to answer. Based on the feedback, we designed this hybrid app (mobile and web) to help novice programmers to post "better" questions on Stack Overflow. The user "scratches" his question in the app and receives suggestions on how to improving the quality of the question. The suggestions are the result of a Natural Language Process (NLP) analysis over the question, which aimed identifying the "good" characteristics that are missing.

---

[1] http://stackoverflow.com/company/about

The survey application resulted in a list with sixteen characteristics. Most respondents agreed that only a little portion of Stack Overflow's questions have high quality. Previous studies have already found that it is possible to teach people to ask better questions (Sullins *et al.*, 2015). Thus, this assistance could be very useful to improving the general question quality in the community.

This work is organized as follows. Section 2 presents related work and how what we are proposing differs from them. Section 3 details the results from the survey application, as well our list of "good characteristics". Section 4 explains how we plan to change the usual Q&A process and Section 5 describes how the app works. In Section 6, we discuss the conclusions and future work.

2 RELATED WORK

According Liu, Bian and Agichtein (2008), low quality questions often lead to bad answers whereas high quality questions usually receive good answers. In (Souza *et al.*, 2016a), it was presented a case study that demonstrated the advantages of writing high quality questions over poor quality questions, in real world situation. In addition, Ravi *et al.*, (2014) analysis showed that higher quality questions continue to garner interest over time in comparison to lower quality questions.

However, writing high quality question may not be intuitive for all. In a CQA, a good question is not just one that is found to be useful by other people: a question is good if it is also presented clearly and shows prior research (Ravi *et al.*, 2014). Hao, Shu and Irawan (2014) identified 14 features that will help in the increasing of the quality of questions, in Stack Overflow. In their work, the authors found that the following content-related features are positively associated with question quality: "w" word, completeness, and subjectivity; while question length, title length, number of tags, code snippet presence, complexity and politeness are negatively associated. Later, a study from (Baltadzhieva and Chrupała, 2015), identified that questions containing incorrect tags, or that are too localized, subjective or off topic are considered of bad quality. On the other hand, the presence of an example has a positive effect on the question score and the number of answers (Baltadzhieva and Chrupała, 2015).

Mamykina and Manoim, (2011) found that 92% of Stack Overflow questions are answered in a median time of 11 minutes. Their research suggests that aspects of the question may influence the speed of response, with, for example, questions that invite discussion being less likely to receive fast responses. According Treude, Barzilay and Storey (2011) the most common use of Stack Overflow is for how-to questions. The site is also effective for code reviews, explaining conceptual issues and answering newcomer questions (Treude, Barzilay and Storey, 2011). The type of question is not the only factor for getting good answers. Other factors seem to include: the technology in question, the identity of the user, the time and day in which the question was asked, whether the question included a code snippet, or the length of the question (Treude, Barzilay and Storey, 2011). The thoughts about the influence of the question length are mixed and contradictory and further research is still necessary in order to provide better insights in its importance for the number of answers and the question score, according (Asaduzzaman *et al.*, 2013).

Asaduzzaman *et al.*, (2013) proposed the taxonomy to explain why questions remain unanswered in Stack Overflow. The top five reasons are: "Fails to attract an expert member", "Too short, unclear, vague or hard to follow", "A duplicate question", "Impatient, irregular or inconsiderate members" and "Too hard, too specific or too time consuming". Understanding those factors that contribute to questions being answered as well as questions remain ignored can help information seekers to increase their chances of getting answers from the community.

Dror, Maarek and Szpektor, (2013) proposed using this information to give users immediate feedback about the ability of his question in attracting answers. Imagine that a user is preparing to broadcast a question in a CQA. If he knew which factors can affect response rate, he could shape his request to fit these factors and theoretically improve his chances of finding help. According Sullins *et al.*, (2015) it is possible to teach people to ask better questions. Their case study revealed that participants in the question training condition asked significantly more "deep" questions on the post-test than did the participants in the control condition.

3 A NEW WAY TO ASK QUESTIONS ONLINE

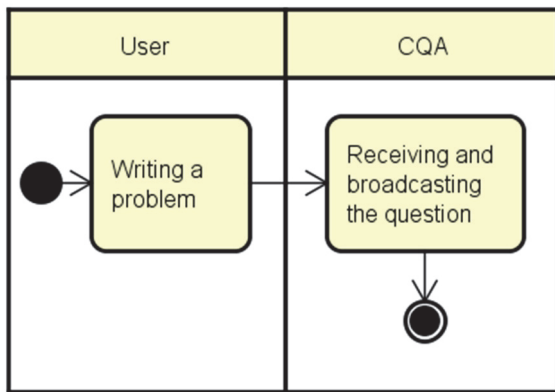Figure 1 illustrates the traditional social query process.

Figure 1: The traditional social query process.

The process starts with the user accessing the website. He, alone, phrases his problem and shares with all users in the collaborative environment. In Figure 7, we used a CQA as example, but it could be any social context like a social network or e-mail's group, for instance. Some collaborative environments could even provide recommendations of which users are able to respond (query routing).

However, the literature review opens the following research opportunity: assisting the user in the task of including these "good" characteristics into the question structure in order to enhance Q&A experience (improving both question quality and attractiveness). We could change the way how social query works.

Basically, we could include a step, before the question is released, in which we assist the user in the task of formulating his problem through the system interface. This assistance aims to help users insert "good" characteristics into their questions. Thus, they are, theoretically, improving their chances of finding help. This new process is presented in Figure 2.

The main difference between the two processes is that, instead of writing the question alone, in this new one, the user is somehow assisted by the website interface. The user interface (UI) "knows" which characteristics a question should have. It analyzes the question and searches for the presence and absence of these "good characteristics".

The UI could give tips on what "good characteristics" are missing in the question's structure or suggest rewritten versions of the original question, but with the "good characteristics" already implemented. For instance, if most questions that are answered have a certain length, the analysis consists in checking if the new question has this exact length
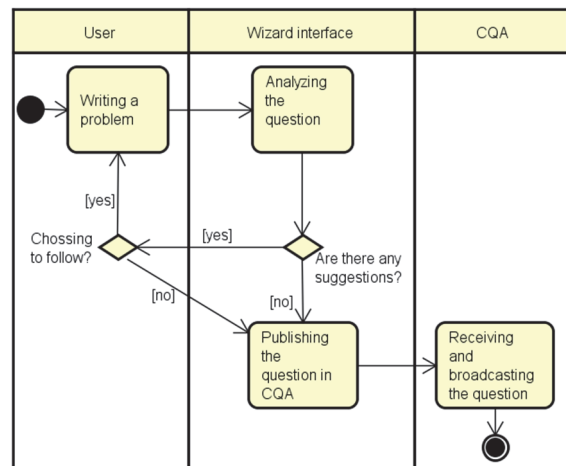

Figure 2: New social query process.

or is close to it. If it does not have it, adjusting the question length will be one of the suggestions outputted by the UI. User will receive this feedback and decide if he wants to follow it or not. He also decides in what extension applying the suggestions. In addition, if we are in a context where query routing works, during "Assistance Phase", the user could also be inquired about reducing the scope of the Expert Search establishing demographic filters. After "Assistance Phase", the query routing would proceed normally.

This "assistance phase" would be more efficient to improve question quality and attractiveness than using the "good characteristics" list as guidelines to users. Imagine a context where there are fifty desirable features that a question could have. It is unlike that all users will study this entire list to ask good questions. However, if the UI "hints" users only with suggestions about features that are missing in the question, probably it would be easier to him to follow.

Including this "assistance phase", however, would demand previously identifying which are the "good characteristics" that a question should have at that context. It is important to highlight that these "good characteristics" are strongly related to the studied context. The list of "good" characteristics is, in Figure 8, obtained through the investigation of CQA's questions history. However, this list of characteristics could also be obtained through: (1) interviews with active users asking them which factors attract them to answer a question; and (2) surveying the literature about question asking in that environment to identify good practices and characteristics that impact response rate. We surveyed the literature to present a preliminary list

of "good characteristics", but, before designing our solution, we also collected users' opinion through a questionnaire, in order to identify which are the "good characteristics" that a question should have to motivate them to answer.

In next section, we discuss these and others outcomes of the survey application.

4 SURVEY APPLICATION

We elaborate a survey to ask users from Programming CQA how close our first drawing of "good characteristics" was from what they was looking for in questions. In addition, we asked which characteristics they were expecting that were not included in the list yet. We collected 400 answers. Before summarizing the results of the survey application, we want to start describing the profile of people who answer the questionnaire.

4.1 Respondents' Profile

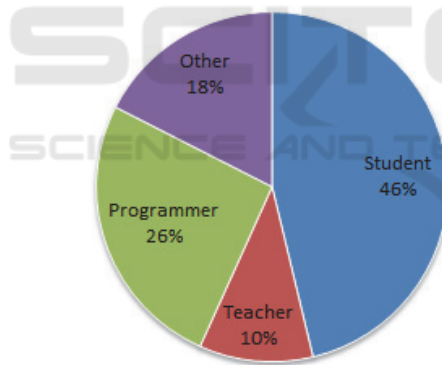Figure 3 describes the occupation of the respondents.

Figure 3: Occupation of the Respondents.

Students represent the larger part of our sample (46%). Since the questionnaire was broadcasted through the university channels, this was expected. The remaining 44% work with programming related jobs (analyst, tester and, of course, programmer). Only 10% of them work as a professor.

Most respondents access these sites in a weekly (39%) or daily (36%) basis. In addition, we also found that while most Programmers (62%) accesses these sites in a daily basis, most Students (41%) are weekly visitors. Probably, this is consequence of their relation with Programming. While professional Programmers deal with Programming daily, as part of their profession, Students check the site

sporadically, since their relationship with Programming is not so intensive, when they are facing homework or just studying.

Figure 4 shows which activities respondents perform when they visit these sites (multiple choices were allowed).
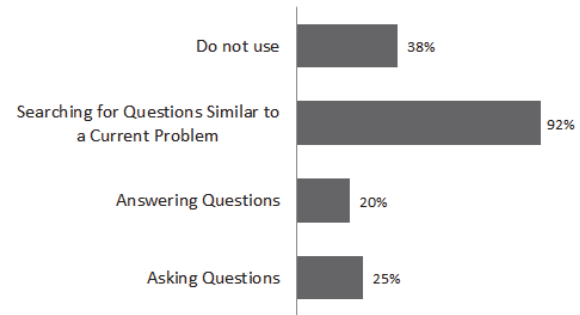

Figure 4: Actions of the Respondents.

We found interesting that *asking questions* and *searching for questions similar to a current problem* were such common roles. People run to Stack Overflow more often than the own documentation of the technology that they are using, when they are facing a problem. This just highlights the collaborative aspect of programming. Unfortunately, our sample is only composed 20% by people who actually answer questions. However, this was already expected, since the respondents are a small portion of the entire community (Furtado *et al.* 2013).

4.2 Respondents' Open Suggestions

Since we did not want to influence respondents with our list, we started asking which characteristics they believed was related with question attractiveness and quality, in their own opinion. We used an open question and we tabulated answers, in order to identify the more frequent suggestions.

Table 1 shows the top five characteristics (grouped by respondent occupation).

In Table 1, we can see that two characteristics were related to both question attractiveness and quality. They are: (1) *objectivity*, and (2) *clarity*. These are all subjective characteristics. It also worth mention that programmers like to answer questions from people who know about the topic they are asking, indicating that programmers do not like to answer newbies' questions. In addition, programmers and professors associated the code or example presence with question quality.

Table 1: Number of mentions of characteristics related to question quality and attractiveness.

Characteristic related with question attractiveness				
Characteristic	**Prof.**	**Stud.**	**Prog.**	**Other**	**Total**
Example or code.	18	77	56	24	175
Objectivity.	17	74	38	35	164
Clarity.	12	57	16	20	105
Short description.	7	30	16	14	67
Coherence between title and description.	5	33	12	14	64
Characteristic related with question quality				
Characteristic	**Prof.**	**Stud.**	**Prog.**	**Other**	**Total**
Clarity	9	61	22	21	113
Objectivity	10	47	26	17	100
Show that knows the topic.	7	32	35	19	93
Example or code.	9	24	30	9	72
Be polite.	5	26	12	11	54

4.3 Respondents' Agreement Rate

In (Souza *et al.*, 2016a), we performed a literature review that aimed to draft a first version of the "good" characteristics list. In (Souza *et al.*, 2016b), we found interesting correlations between the presence of these "good" characteristics and questions' performance.

We used the questionnaire to capture people's opinion about this version of our "good" characteristics list. In the last part of the questionnaire, we presented this first drawing and we asked if the respondent agreed that the presence of each characteristic was important. Table 2 shows the percentage agreement of the respondents with each "good" characteristic.

In Table II, we can see that two characteristics have a high disagreement percentage: *Title entirely written in capital letters* and *prioritizing long description*. We believe that *long description* requires more effort from the user to answer, discouraging most people. In addition, the disagreement related to the use of *capital letters* is due its Internet meaning, which it is usually taken as yelling. These characteristics were not considered later on our study.

Table 2: Agreement rate with "suggested" characteristics.

Characteristic	Yes	No	None
Well-chosen title	**94%**	1%	5%
Title partially written in capital letters	10%	34%	**56%**
Title entirely written in capital letters	2%	**62%**	36%
Coherence between question description and title	**97%**	1%	2%
Understandable description	**95%**	1%	4%
Including a vocative	12%	30%	**59%**
Prioritizing short description	**47%**	18%	36%
Prioritizing medium size description	43%	6%	**51%**
Prioritizing long description	12%	**55%**	34%
Showing an example	**82%**	5%	13%
Avoiding a large amount of code	**51%**	22%	28%
Avoiding description with only code	**60%**	15%	25%
Restricting each question to a single problem	**72%**	7%	21%
Including greetings	42%	10%	**48%**
Using proper language	44%	9%	**47%**
Avoiding creating duplicate questions	**81%**	4%	16%
Avoiding creating factoid questions	27%	31%	**42%**
Do not create homework questions	**76%**	5%	19%
Including links related to the question	**65%**	5%	30%
Combining links with partial content	**53%**	7%	40%

The following characteristics received a high "indifferent" rate: *Title partially written in capital letters, including vocative, prioritizing medium size description*, *including greetings*, *using proper language*, *avoiding creating factoid questions* and *combining links with partial content*. These characteristics were not entirely considered on our study, as we will explain next.

After analyzing the responses for both, the open and the objective, questions, we reduce our list of "good" characteristics to the following:

1) **Objectivity**: objectivity was the top characteristic for both quality and attractiveness,

according respondents. However, it is a strongly subjective concept to identify. Most dictionary definitions emphasize the shortness aspect of objectivity. Thus, we will reduce Objectivity to these other characteristics: (1) *Restricting each question to a single problem*; and (2) *Prioritizing short description*. Although we summarized objectivity in only these two features, we believe that the objectivity's presence is also related to other features included on the list, such as "Clarity" and "Well-written description".

2) Clarity: clarity was another top mentioned characteristic and, similarly to objectivity, it is a strongly subjective concept too. We define clarity as the quality of being easily understood. We decide for checking clarity through these three other characteristics: (3) *Coherence between question description and title*; and (4) *Making the problem the more evident as possible in the description*.

3) Well-written description: although users want to help, poorly written questions (vague or incomplete, for instance) will discourage them. For this reason, it is important to make an effort in writing self-contained questions. We believe that these hints would help in this matter: (5) *Including example or code*; (6) *Including links related to the question*; (7) *Combining links with partial content*; and (8) *Avoiding description with only (or a large amount of) code*.

4) Be polite: – there were (a few) mentions, in open suggestions, that the politeness of the asker was one of the factors that courage people into helping. Thus, we included in the final list these etiquette rules for asking questions online: (9) *Avoiding creating duplicate questions*; (10) *Do not create homework questions*; (11) *Including greetings*; and (12) *Using proper language*.

This list summarizes what the community wishes all questions look like. Thus, based on this list, we develop a *suggestion engine*, which analyzes questions looking for the absence of each one of these 16 characteristics and giving feedback to the user on how to improve his question's quality by adding the ones that are missing. The NLP techniques were programmed using CoGroo (http://cogroo.sourceforge.net/), OpenNLP (https://opennlp.apache.org/), and LanguageTool (https://www.languagetool.org/).

Next, we will present the mobile app, which it is an instance of this *suggestion engine*.

5 PRESENTING THE APP

We named the app Question's Advisor, due his role of helping but without forcing the user to follow it. The software was developed using Progressive Web Apps (PWA) technology, which makes it suitable for Desktop and Mobile clients. One can access our prototype using this address http://appif.herokuapp.com/.

In Figure 5 (left), we show the login screen of the Question's Advisor (we will illustrate using the mobile vision). When the user clicks in "authorize", he is requested to login using his Stack Overflow credentials (center). This process is mandatory to allow the app publishing questions in Stack Overflow using the user's account.
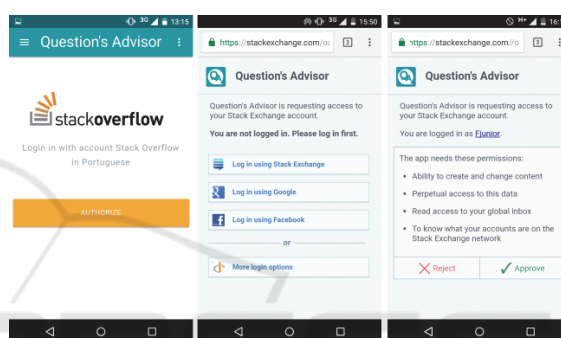

Figure 5: Login screen and authorize screen from question's advisor.

After login, we ask the user the mandatory permissions for reading and writing using his account, this also can be seen in Figure 5 (right). After, the user is directed to the home screen of the app, that it is presented in Figure 6.
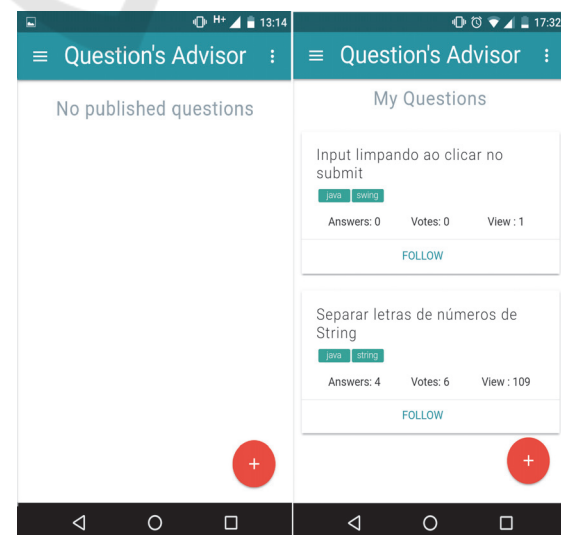

Figure 6: Home screen from Questions' Advisor.

In Figure 6, we can see an empty list (left), but the home screen is also able to list all questions published by the user (right). Below each question on the list, he has a summary of the question's performance regarding number of answers, points and views. When clicking in one of questions, the user is directed straight to Stack Overflow, where he can check the activity over his question. The "plus" button in the home screen (or the lateral menu) directs the user to the new question screen, inside the app. This screen is presented in Figure 7.
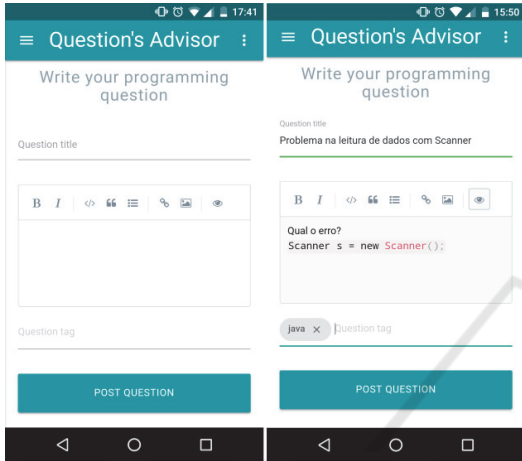

Figure 7: Writing a new question using Question's Advisor.

In Figure 7, we can see that the description of the question may include code snippets and they will appear with a different style to make the reading easier. Moreover, the user can also add tags to describe the question subject. These tags will be published on Stack Overflow too. While making usability tests, most users prefer writing their question accessing the app through a computer instead a smartphone, but both "visions" offer the exact same functionalities.

After writing the question, the user will click on "publish". However, before the sharing the question, he will receive a feedback from the app about how he can improve the question; this can be seeing on Figure 8 (left).

In Figure 8 (center), we can see the list of suggestions too. The suggestions are the result of a Natural Language Process (NLP) analysis over the question, which aimed identifying the "good" characteristics that are missing. The user can roll down this list and he will see the "ignore" button and the "review" button. If the user clicks in "ignore", the question will be immediately shared on Stack Overflow. If the user clicks in "review", he will have the chance of following the suggestions and rewriting the question.
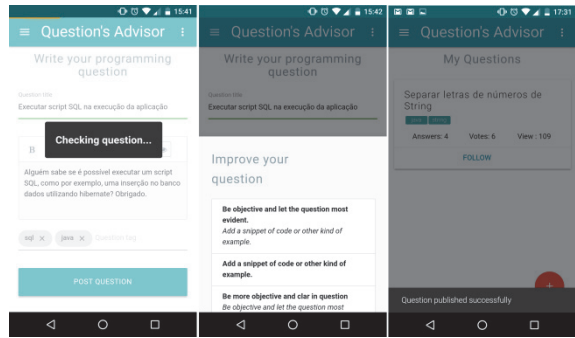

Figure 8: Checking the Question and Suggesting Improvements.

After editing the question, the user will click again in "publish" and the Question's Advisor will calculate if the number of "good" characteristics has increased in comparison with the last analysis (indicating a gain in the question quality). In the case where the number of "good characteristics" fell, the user will keep receiving suggestions on how to improve the question (he will be able to ignore again, if he prefers).

After publishing, the user will receive a confirmation that his question was published in the Stack Overflow, as presented in Figure 8 too (right), and the question will appear in the Home Screen.

6 CONCLUSIONS AND FUTURE WORK

Programmers often recur for online communities in order to find help for a current problem that they are facing. However, CQAs like Stack Overflow, while are efficient for this goal, does not ensure an answer for all question. The number of ignored questions is constantly increasing. One of the reasons why questions remain unanswered is due their low quality. In addition, some studies found a correlation between the question's characteristics and its ability of drawing attention and being answered.

In this work, we conducted a mixed study aiming identifying what programmers are looking in a question that they decide to answer. We designed a questionnaire where people could suggest their own thoughts and also agree or disagree with our opinion. Based on these answers, we developed a tool to help programmers to write high quality questions. Our solution analyses the original question written by the user and suggests including the missing "good" characteristics. We named the app Questions' Advisor and it is available for desktop and mobile

clients. We believe it could be very helpful, especially for novice programmers.

For future work, we aim to test our solution locally, collecting feedback, and reporting what happened. In addition, we need to do a further investigation on how the presence of these features relates to question's performance. This research would also have the goal of finding new relevant characteristics that were not suggested though the questionnaire. Last, we want to design a plugin that allows people to add our suggestion engine in any text area component. This way, our solution could be added into small learning communities from Moodle or Facebook, for instance.

ACKNOWLEDGEMENTS

REFERENCES

Mamykina, L., Manoim, B., Mittal, M., Hripcsak, G., and Hartmann, B., 2011. Design lessons from the fastest q&a site in the west. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 2857-2866).

Hao, G. K. W., Shu, Z., and Irawan, J., 2014. Good or Bad Question? A Study of Programming CQA in Stack Overflow.

Baltadzhieva, A., Chrupała, G., 2015. Question quality in community question answering forums: a survey. *Acm Sigkdd Explorations Newsletter*, vol. 17(1), 8-13.

Asaduzzaman, M., Mashiyat, A. S., Roy, C. K., and Schneider, K. A., 2013. Answering questions about unanswered questions of stack overflow. In *Mining Software Repositories* (MSR), 2013 10th IEEE Working Conference (pp. 97-100).

Teevan, J., Morris, M. R., and Panovich, K., 2011. Factors Affecting Response Quantity, Quality, and Speed for Questions Asked Via Social Network Status Messages. In *International Conference of Webblogs and Social Media (ICWSM)* (pp. 630–633).

Sullins, J., McNamara, D. S., Acuff, S., Neely, D., Hildebrand, E., Stewart, G., and Hu, X., 2015. Are You Asking the Right Questions: The Use of Animated Agents to Teach Learners to Become Better Question Askers. In *The Twenty-Eighth International Flairs Conference* (pp. 479-482).

Liu, Y., Bian, J., Agichtein, E., 2008. Predicting information seeker satisfaction in community question answering. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 483-490).

Souza, C., Aragão, F., Remígio, J., Costa, E., & Fechine, J., 2016a. Using CQA History to Improve Q&A Experience. In *International Conference on Computational Science and Its Applications (ICCSA)* (pp. 570-580). Springer International Publishing.

Ravi, S., Pang, B., Rastogi, V., and Kumar, R., 2014. Great Question! Question Quality in Community Q&A. In *International AAAI Conference on Weblogs and Social Media (ICWSM)* (pp. 426-435).

Treude, C., Barzilay, O., and Storey, M. A., 2011. How do programmers ask and answer questions on the web?: Nier track. *In Software Engineering (ICSE), 2011 33rd International Conference* (pp. 804-807). IEEE.

Dror, G., Maarek, Y., and Szpektor, I., 2013. Will my question be answered? predicting "question answerability" in community question-answering sites. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 499-514). Springer Berlin Heidelberg.

Furtado, A., Andrade, N., Oliveira, N., and Brasileiro, F., 2013. Contributor profiles, their dynamics, and their importance in five q&a sites. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (pp. 1237–1252).

Souza, C., Remígio, J., Aragão, F., Costa, E., and Fechine, J., 2016b. Investigating How "Good" Characteristics' Presence Are Related with Questions' Performance: An Empirical Study on a Programming Community. In *Intelligent Systems (BRACIS), 2016 5th Brazilian Conference* (pp. 289-294). IEEE.