Investigating How "Good" Characteristics' Presence are Related with Questions' Performance

an Empirical Study on a Programming Community

Cleyton Souza, José Remígio, Franck Aragão
Department of Systems Analysis and Development
Federal Institute of Education, Science and Technology (IFPB)
Monteiro, Brazil
{cleyton.caetano.souza, fjunior.aragao, jose.rafael.feitosa}@gmail.com

Evandro Costa, Joseana Fechine
Laboratory of Artificial Intelligence (LIA)
Federal University of Campina Grande (UFCG)
Campina Grande, Brazil
{ebcosta, joseanamf}@gmail.com

***Abstract*—Social query is the practice of sharing questions through collaborative environments. In order to receive help, askers usually broadcast their requests to the entire community. However, the prerequisite to receive help is to have the problem noticed by someone who is able and available to answer. Previous works have identified a correlation between the characteristics of the questions and the outcome of receiving or not an answer. These findings suggest that there are some characteristics that are more likely to attract the helpers´ attention. Our proposal is to analyze the CQA history to identify how these so-called "good" characteristics affect the performance of shared questions. Our results suggest that: (1) answered questions present more of these "good" characteristics than unanswered ones; (2) the more "good" characteristics are present in a question, the more people it attracts; (3) the more people are attracted by a question, the faster it will be answered and more responses will be received; (4) answered questions attract more people than unanswered ones; (5) difficulty does not play a role in attracting people.**

Keywords—Social Query; Community Question Answering Sites; Question Quality; Question Redesign; Question Attractiviness.

I. INTRODUCTION

The practice of sharing questions through social media is known as social query. Community Question and Answering sites (CQA) are collaborative environments entirely dedicated to asking and answering questions practice [1]. The most common sharing strategy is broadcasting the problem to everyone in the community. There is two possibilities to receive help: (1) someone who is able notices the question and decides to answer or (2) someone who may know other who probably is able to help decides to forward the question [2]. However, there is no guarantee if these situations will happen neither when.

Some studies found a correlation between the characteristics of the question and the outcome of receiving or not receiving answers [8, 9, 10]. While analyzing answered questions history, it is possible to identify common characteristics among answered questions. We believe that if users knew which characteristics attract others' attention, they could use this piece of information to improve their chances of finding help.

Thus, in this work, we investigate how the presence of certain characteristics affects the performance of questions shared through CQA. Our goal is to test the following hypotheses:

- H_1: *Answered questions have more "good" characteristics than unanswered ones*.
- H_2: *There is positive correlation between the number of "good" characteristics and the number of people attracted by the question*.
- H_3: *There is a negative correlation between the number of "good" characteristics and the time to first response*.

To test our hypotheses, we gathered questions from a programming CQA and described them by the presence and absence of these "good" characteristics. Our results showed that answered questions present more of these characteristics than questions that remain unanswered. We also found a weak correlation between the presence of these "good" characteristics and the number of people that the question attracts. However, we could not statistically confirm a direct correlation between the time to first response and the number of "good" characteristics. In addition, we tested correlations involving the number of people attracted and the answered questions, time to first response, number of responses and the question difficulty.

The remainder of this paper is organized as follows: Section 2 presents related work; in Section 3, GUJ is presented (the context of our study) and the characteristics which are considered "good" are detailed; the evaluation is detailed in Section 4 (including methodology, results, discussion and threats to validity); Section 5 presents how the results could be used in real world to improve Q&A experience; and, finally, Section 6 ends with conclusions and future work.

II. Related Work

The habit of sharing questions on the Web was born on newsgroup during the early years of the Internet, but it was "professionalized" by CQAs [6]. In CQAs, users voluntarily post and answer questions published by other users aiming mutually to help each other [1]. According to Furtado and Andrade [7], one of the main issues of these communities is that the experts (user that frequently offers great answers) are the ones that interact less. Thus, it is mandatory to keep them engaged with the CQA with high quality questions.

According to Burke et al. [6], in Usenet groups, introductions referencing lurking and a personal connection to the topic of discussion increase the likelihood of getting a reply. In Yahoo! Answers[1], Yang et al. [9] found that medium length questions are less likely to get answered, as well as questions from "other" category or with low similarity with their assigned category. This happens because these questions fail to attract attention of most users. Asaduzzaman et al. [10] conducted a qualitative analysis to understand why questions remain unanswered in Stack Overflow[2] services. The top five reasons are:(1)"Fails to attract an expert member"; (2) "Too short, unclear, vague or hard to follow"; (3)"A duplicate question"; (4) "Impatient, irregular or inconsiderate members"; and (5) "Too hard, too specific or too time consuming".

Regarding personal social networks' studies, Teevan, Morris and Panovich [5] found that, on Facebook, a concise style of question-asking, a predefined audience, and the inclusion of a question mark were associated with more and higher quality responses within shorter periods of time. In [11], they also found that young people and people with larger social networks are more likely to receive answers. In addition, they established a correlation between the length of the questions and the receiving of response: questions with extra sentence are less likely to receive "yes/no" answers or requests for clarification. Lampe et al. [12] found that the question type affects the performance of questions shared on Facebook. Recommendation posts receive more responses than any other question type, while favor requests usually take a long time until they receive a first response. Comarela et al. [13] conducted a study to understand the factors that affect response rate in Twitter and found that tweets with hashtags and URL are more likely to be retweeted, and tweets with mentions are more likely to receive a reply. This last result supports Nichols and Kang's claim that directing questions is more effective than broadcasting to receive answers.

All these findings could be used to improve the likelihood of one getting answers [3]. Imagine that a user is preparing to broadcast a question in a social context. If he had this knowledge about which factors can affect response rate, he could shape his request to fit these factors and theoretically improve his chances of finding help [3]. Moreover, this could be used to improve the quality of questions and, consequently, of the answers [14]. The goal of teaching students to ask better questions was explored by Sullins et al. [15]. Their results revealed a significant change in the quality of questions generated on the post-test as a function of condition (participants in the question training condition asked significantly more "deep" questions on the post-test than did the participants in the control condition).

These results open interesting research opportunities like if it is possible to improve Q&A experience through the investigation of CQA history. Through the analysis of question history, common characteristics among answered questions could be identified. While users are phrasing new questions, we could suggest that they add these characteristics into their request, improving both question quality and question attractiveness. And, finally, these redesigned questions could be easier to respond, whether just for being clearer or for having a specific characteristic that attracts the other's attention.

Dror, Maarek and Szpektor[3] propose to provide a type of "heads up" to askers by predicting how many answers, if at all, they will get. According to them, giving a preemptive warning to the asker at posting time should reduce the frustration effect and hopefully allows askers to rephrase their questions if needed. Souza et al. [8] present a mobile app that allows users to share questions with their friends on Facebook. Through this app, users inform their questions and receive suggestions to increase the probability of receiving answers. The suggestions range from tips to rephrase the question and to indications about who probably knows the answer (a person or a demographic group).

In this work, we are investigating the relationship between questions' characteristics and their performance after shared. We believe this is an important step in order to understand if it is possible to increase the likelihood of receiving answers by assisting the user in the task of rephrasing his question.

III. Investigation CQA History

We developed our study in a Brazilian CQA about programming called GUJ[3]. This is the largest programming community in Brazil, with almost 200 thousand users. GUJ means Java User Group, in Brazilian Portuguese. The website was created in 2001 and it works like a forum. Users access GUJ and publish questions like a new thread. When other users access GUJ, they are presented to a list of threads ordered by the last interaction time. They can access the thread and reply to its author. Since its beginning, it has been made more than 300 thousand questions and it has been exchanged almost 2 million messages.

We conducted a qualitative study in order to identify which are the most common characteristics in the answered questions. We started gathering a sample of threads from GUJ containing both answered and unanswered questions. This study included the analysis of this sample. We described questions using attributes like: question length, title length, question and title coherency, code presence, greetings presence, question topic, difficulty level, etc. After outlining a list of characteristics, we carried out a literature review searching for articles about asking good programming questions. Since GUJ is strongly popular among Brazilian students, there is a lot of material to

[1]http://answers.yahoo.com

[2]http://stackoverflow.com/

[3] http://www.guj.com.br

help newcomers to ask "good" questions. We confronted and combined both analyses and they resulted in the following list of characteristics which a question should have to attract more responses.

Title related characteristics – The title is the first contact of potential responders with the question. The title should be a summary of the problem and it cannot be too short or too long. Regarding a good title characteristic, users should prioritize: understandable title; medium size title; title with capital letters; and a title coherent with the subject of the question description.

Description related characteristics – After the user has been attracted by the title, the potential helper will read the problem description. We observed that some users do not want to "waste their time" looking at a long code or following a link. Thus, questioners should keep the description with enough information that anyone can answer without additional reading effort. However, we are aware that being concise and clear is not always an easy task. In regard to good description characteristics, users should prioritize: using understandable description; including a vocative; avoiding too long or too short description; showing an example, but avoiding a large amount of code; avoiding description with code only; and, when including links, combining them with partial content.

Behavior related characteristics – Helpers will be more willing to answer questions from "good" users. We identify that users who follow a community normative sense have more chances of receiving answers. These "good" users are relatively polite, grateful for receiving help, and aware of a correct way to behave that is not written anywhere, but it is unconsciously followed. Regarding this matter, we identify the following good practices: restricting each question to a single problem; including greetings; obviating demanding language; using proper language; avoiding creating duplicate questions; avoiding creating factoid questions (this kind of problem is well solved through search engine use); and do not create homework questions (this can be misunderstood as laziness).

Since this list of characteristics emerged from a literature review, we are assuming that they are, at least, good characteristics that questions should have. On the other hand, the fact that their presence is related to question attractiveness and responsiveness will be quantitatively verified in the next section.

IV. Evaluation

In this section, we describe our empirical study. First, we present the methodology. Next, we show our results, followed by discussion. Finally, we discuss the limitations of our work.

A. Methodology

To test our hypotheses, we gathered questions from GUJ and described them by the presence and absence of "good" characteristics. All data used in this work is available for download[4].

[4] Data is available at this link https://drive.google.com/open?id=0B0jcFw9ETL-xeHhsVmRQeGxTbTg.

Questions were randomly selected using GUJ's search tool. Since we always chose top retrieved results, few outliers were part of the answered questions sample (like one question that attracted 3,500 page views). Although these questions did not represent well the majority of questions shared in GUJ, we decided not to interfere in which questions would be part of the sample.

We separated retrieved questions into two samples: answered questions and unanswered questions (with a hundred questions each). We compared the performance of these questions regarding: time to first response, number of attracted people (based on the number of page views) and number of received answers. Since GUJ's architecture does not show the number of people attracted by unanswered questions, we had to answer those with a fake response, but we deleted the answer, just after collecting the information.

B. Results

First, we want to show data about the questions that compose our samples. Figure 1 presents how questions are distributed through a three-class range of difficulty (easy, medium and hard – these labels were assigned for two computing students).

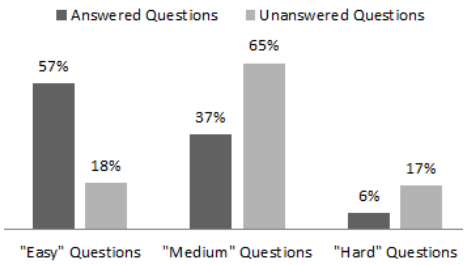

Fig. 1. Difficulty range of the questions

In Figure 2, we show the boxplot for the number of page views for easy questions and for not easy questions (medium and hard).We excluded outliers just for a better visualization of data, but they will be part of our analysis.

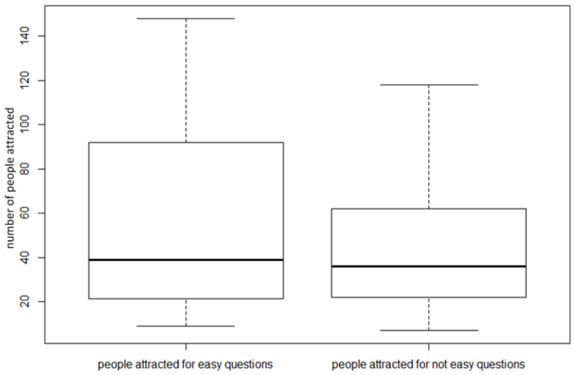

Fig. 2. Boxplot about the number of people attracted by easy and not easy questions

Table 1 shows the percentage of questions in which each "good" characteristic, detailed in the last section, was present.

TABLE I. DATA FROM THE SAMPLES OF ANSWERED AND UNANSWERED QUESTIONS

Characteristcs	% presence in answered questions' sample	% presence in unanswered questions' sample
Understandable title	79%	56%
Medium size title ($2 < words < 6$)	77%	77%
Title with capital letters only	0%	0%
Title with capital letters partially	10%	2%
Coherency between question description and title	99%	86%
Understandable description	92%	56%
Including a vocative	45%	63%
Avoiding too short description	85%	74%
Avoiding too long description	98%	94%
Showing an example	67%	59%
Avoiding a large amount of code	99%	83%
Avoiding description with code only	100%	99%
Restricting each question to a single problem	96%	94%
Including greetings	41%	44%
Obviating demanding language	100%	91%
Using proper language	98%	89%
Avoiding creating duplicate questions	100%	89%
Avoiding creating factoid questions	100%	100%
Do not create homework questions	97%	93%

In Table 1, we observe that only a small portion of questions used the capital letters trick to attract attention. In addition, in our sample, there are no questions with links, thus, we did not display this feature in Table 1.

Next, in Table 2, we show the performance of questions regarding the number of attracted people (mean, median, min and max). In addition, in the case of answered questions, we also present the mean time to first response and the mean number of responses.

TABLE II. QUESTIONS' PERFORMANCE

Characteristic	All Questions	Answered Questions	Unanswered Questions
Mean of attractedpeople	103	175	31
Median of attractedpeople	37	71	26
Max of attractedpeople	3500	3500	96
Min of attractedpeople	7	10	7
Mean time to first response	-	193 min	-
Mean of responses	-	6	-

In Table 2, we observe that there is a discrepancy between the mean and the max of people attracted by the sample of answered questions. Analyzing our data, we found four answered questions with more than a thousand views. This is explained by the nature of these questions, as they are about common doubts of new programmers. For instance, the question that had 3,500 views is about "How to convert a String in a number".

We used *Shapiro Wilk* test to verify the normality of our data. All tests rejected the null hypothesis, which means that no distribution is normal. Thus, we used *Spearman's Rank Correlation Coefficient* to establish the statistical dependence among variables. Table 3 shows the results of the correlation tests.

TABLE III. RESULTS OF CORRELATION TESTS

Variables	p-value	Spearman rho
Percentage of "good" characteristics in ALL QUESTIONS and Number of Attracted People	0.0122*	+0.176914
Percentage of "good" characteristics in UNANSWERED QUESTIONS and Number of Attracted People	0.8639	+0.017356
Percentage of "good" characteristics in ANSWERED QUESTIONS and Number of Attracted People	0.6343	−0.048152
Percentage of "good" characteristics in ANSWERED QUESTIONS and Time to First Response	0.2321	+0.120578
Percentage of "good" characteristics in ANSWERED QUESTIONS and Number of Users who Responded	0.9883	+0.001483
Percentage of "good" characteristics in ANSWERED QUESTIONS and Number of Responses	0.8286	−0.021920
Number of Attracted People and Time to First Response	0.0015*	−0.240183
Number of Attracted People and Number of Received Responses	$2.37e-06$*	0.451859

* The tests presented statistical significance (α=5%)

As we can see in Table 3, most correlations did not show statistical significance. Nevertheless, those that showed it, just a week correlation was pointed out.

Next, we will fully discuss all these results.

C. Discussion

Our first hypothesis was H_1:*Answered questions have more "good" characteristics than unanswered questions*. Table 1 summarized the percentage of questions in which each "good" characteristic, detailed in the last section, was present. Nonetheless, to verify this hypothesis, we compared the percentage distribution of "good" characteristics present in each question that belongs to the answered and unanswered sample, respectively. We used *Wilcoxon Signed Rank Test* to compare if the percentage distribution of answered questions was greater than that of unanswered questions. The result was a p-value equals to 1.11e-06. Since the p-value is lower than the significance level (α=5%), we denied the null hypothesis, thus confirming that H_1 is true.

The second hypothesis was H_2: *There is positive correlation between the number of "good" characteristics and the number of people attracted by the question*. Table 3 showed that we used *Spearman's Rank Correlation Coefficient* to establish the statistical dependence among these variables. We

observed that, only when combining both samples, it was possible to confirm a weak positive correlation with statistical significance (p-value=0.0122). The higher the percentage of the characteristics is, the more people will be attracted by the question. These results would be enough to accept H_2. However, when we split questions in their respectively groups, there was no evidence of this weak correlation, as also presented in Table 3. Thus, although we were able to statistically confirm a weak positive correlation among these variables, we believe that a conclusive study would demand a larger sample.

Regarding H_3: *There is a negative correlation between the number of "good" characteristics and the time to first response*. It could not be statistically confirmed, since the p-value (0.2321) was higher than the significance level (α=5%). However, while analyzing our data, we observed that the more people are attracted by a question, the faster it will receive an answer. This was statistically verified (p-value=0.0015), as it can be seen in Table 3. This makes sense, because if more people see the question, the likelihood of receiving an answer will be higher. Since H_2 was also accepted, we have an indicator that the more "good" characteristics a question has, the faster it will get the first response. This contradicts the rejection of H_3. Thus, further analysis is necessary to definitely deny this hypothesis.

Besides testing our three hypotheses and identifying that more page views reduce the time to first response, we also noticed that answered questions naturally attract more people than unanswered ones. We used *Wilcoxon Signed Rank Test* to confirm that the distribution of page views of answered questions was higher than the unanswered ones (p-value= 9.536e-14). We believe that this has, at least, two explanations. First, due to GUJ's architecture, threads with recent interaction are positioned in the top of the feed of all users, thus attracting more attention. The second explanation is that these questions keep attracting people who are facing the same problem, therefore increasing the number of page views. In addition, more people present a weak positive correlation with the number of received responses, as we saw in Table 3.

Finally, we also tested if the question difficulty is associated with its attractiveness. The aim of this analysis was to try to exclude the difficulty as a factor that influences our results regarding attractiveness. According to Asaduzzaman et al. [10], in Stack Overflow, a context similar to ours, hard questions are unlike to receive answers. For this reason, we showed in Figure 2 the distribution of questions of our samples. Although we have a few more easy and hard questions, respectively, in the answered questions sample and in the unanswered questions sample, the overall distribution is balanced. In regard to the idea that easy questions may attract more people, we used *Wilcoxon Signed Rank Test* to verify if more people were attracted by easy questions than by not easy questions (medium and hard). The test resulted in a p-value of 0.1885, which is insufficient to discredit the null hypothesis. This means that there is no statistical difference between the number of people attracted by easy and not easy questions, at least in the GUJ context. This can also be seen graphically in Figure 2 (there is a big intersection between both boxplots).

D. Threats to Validity

This section briefly discusses the limitations in our work. To analyze the validity of the results, we consider the four kinds of threats: external, internal, conclusion and construction. The external validity is related to the approximated truth of conclusions and generalization to the real world; the internal validity corresponds to checking if the results are a consequence of the manipulation that was done and not regarding any other factors; the conclusion validity refers to the correct correlation between what was verified (measured) and the reached conclusions; and, finally, construction validity makes reference to problems in the design and control of the experiment.

In relation to the external validity, a threat to the reached conclusions is that we tested our claims in the Brazilian CQA context about programming. Although, we are confident that our conclusions fit any collaborative environment, we cannot guarantee that the same observations will happen outside GUJ. Another external threat is the size of our sample. However, we supported our conclusions in statistical tests that consider the population size in order to compute the p-value.

With reference to the internal validity, there is the possibility that disregarded random factors may affect our observations. For example, if our definition of an easy question is not accurate, our conclusions regarding this matter are invalid. We asked computing student in the mid-course to evaluate the difficulty of questions. They are intermediate programmers highly experienced in Java Language and we reviewed and discussed this analysis with them. Thus, we believe this is very unlikely. We did ignore the publishing time of questions in our analysis, because when we summarized our data, we found a wide diversity, as expected. Although we had tested, we were unable to identify which role publishing time plays, if any. In addition, we were very careful about building and describing our data, thus reducing the chances of errors.

With regard to the conclusion validity, problems may occur on the findings if late interactions would come or if we established the wrong "good" characteristics. However, we believe that this is also very unlikely, since the flow of new threads on GUJ is intensive. As time goes by, a question thread has lesser chances to receive more replies; and, regarding the set of "good" characteristics, it was based on a mixed study involving literature review and classification task.

In relation to the construction validity, we used questions retrieved by using GUJ' search tool; thus, there is the possibility the questions that compose our data are biased by the search algorithm. However, we believe this is not the case, due to the balance of our sample regarding the number of attracted people, the number of received responses, time to first response and question difficulty. We built a representative sample of the universe of questions. In addition, we were careful to choose the appropriate statistical test to verify our hypotheses. These tests are even resistant to the interference of outliers.

V. Conclusion and Future Work

In this work, we presented a study about how the characteristics of the questions are related to their

attractiveness. We carried out a mixed study to identify which characteristics a question about programming should have. We gathered a sample with answered and unanswered questions from a Brazilian CQA about programming called GUJ. Our aim was to investigate the relationship between the characteristics of the questions and their performance regarding the number of attracted people, time to first response and number of received answers.

Our empirical analysis showed that answered questions present more of these "good" characteristics than unanswered ones. Additionally, we observed that the more "good" characteristics a question has, the more people it will attract; and the more people are attracted by a question, the faster it will be answered and more responses it will get. We also found that answered questions attract more people than unanswered ones, whereas the difficulty does not play a role in attracting people, at least in GUJ's context.

Based on our analysis, the more "good" characteristics a question has, the more people it attracts; this could reduce the time to first response. These results open 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 by including 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. Including this "assistance phase" would demand previously identifying which are the "good" characteristics that a question should have at that context. The list of "good" characteristics could be obtained through: (1) the investigation of CQA's history, (2) interviews with active users asking them which factors attract them to answer a question; and (3) surveying the literature about question asking in that environment to identify good practices and characteristics that impact response rate. The UI could give tips on what features the user should add to 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 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 how to apply the suggestions.

As for future work, we want to investigate the individual impact of each "good" characteristic in the receiving of answers and how the characteristics are related to responses' quality. We would also like to conduct this study in other contexts, like Stack Overflow, and compare results. Finally, we want to apply these results to develop an assistive UI that helps users to phrase their problems including these "good" characteristics; analyzing if this actually enhances Q&A experience (improving both question quality and attractiveness), as we suppose.

ACKNOWLEDGMENT

We want to thank the IFPB – Campus Monteiro for the outstanding support provided during the development of this study.

REFERENCES

[1] Raban, D., & Harper, F. (2008). Motivations for answering questions online. *New Media and Innovative Technologies*.

[2] Souza, C. C. De, Magalhães, J. J. De, Costa, E. B. De, & Fechine, J. M. (2012). Predicting Potential Responders in Twitter : A Query Routing Algorithm. In *Proceedings of the 12th International Conference on Computational Science and Its Applications* (pp. 714–729).

[3] Dror, G., Maarek, Y., & Szpektor, I. (2013). Will My Question Be Answered? Predicting "Question Answerability" in Community Question-Answering Sites. In *Machine Learning and Knowledge Discovery in Databases* (pp. 499–514).

[4] Comarela, G., Crovella, M., Almeida, V., & Benevenuto, F. (2012). Understanding Factors that Affect Response Rates in Twitter. In *Proceedings of the ACM SIGWEB Conference on Hypertext and Social Media (HT'12)* (pp. 123–132).

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

[6] Burke, M., Joyce, E., Kim, T., Anand, V., & Kraut, R. (2007). Introductions and requests: Rhetorical strategies that elicit response in online communities. *Proceedings of the 3rd Communities and Technologies Conference, C and T 2007*, (pp. 21–39).

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

[8] Souza, C.; Magalhães, J.; Costa, E.; Fechine, J.; Reis, R. 2014. Enhancing the Status Message Question Asking Process on Facebook. *International Conference on Computational Science and Its Applications*, (pp. 682-69).

[9] Yang, L., Bao, S., Lin, Q., Wu, X., Han, D., Su, Z., & Yu, Y. 2011. Analyzing and Predicting Not-Answered Questions in Community-Based Question Answering Services. *Twenty-Fifth AAAI Conference on Artificial Intelligence*, (pp. 1273–1278).

[10] Asaduzzaman, M., Mashiyat, A. S., Roy, C. K., & Schneider, K. A. 2013. Answering questions about unanswered questions of Stack Overflow. *2013 10th Working Conference on Mining Software Repositories (MSR)*, (pp. 97–100).

[11] Teevan, J., Morris, M. R., & Panovich, K. 2013. Does anyone know how to get good answers? How social network questions shape replies. *Technical Report.* doi:MSR-TR-2013-62

[12] Lampe, C., Gray, R., Fiore, A., & Ellison, N. 2014. Help is on the way: patterns of responses to resource requests on facebook. In *Proceedings of the 2014 conference on Computer supported cooperative work - CSCW '14*,(pp. 3–15).

[13] Comarela, G., Crovella, M., Almeida, V., & Benevenuto, F. 2012. Understanding Factors that Affect Response Rates in Twitter. In *Proceedings of the ACM SIGWEB Conference on Hypertext and Social Media (HT'12)*, (pp. 123–132).

[14] Baltadzhieva, A., & Chrupała, G. 2015. Question Quality in Community Question Answering Forums. *ACM SIGKDD Explorations Newsletter*, (pp. 8–13).

[15] Sullins, J., Mcnamara, D., Acuff, S., Neely, D., Hildebrand, E., Stewart, G., & Hu, X. 2015. Are You Asking the Right Questions : The Use of Animated Agents to Teach Learners to Become Better Question Askers. In *Proceedings of The Twenty-Eighth International Flairs Conference*, (pp. 479–481).