

Using CQA History to Improve Q&A Experience

Cleyton Souza^{1,2}(✉), Franck Aragão¹, José Remígio¹, Evandro Costa³,
and Joseana Fechine²

¹ Federal Institute of Education, Science and Technology of Paraíba - IFPB,
Monteiro, Paraíba, Brazil

`cleyton.caetano.souza@gmail.com`

² Federal University of Campina Grande - UFCG, Campina Grande, Paraíba, Brazil

³ Federal University of Alagoas - UFAL, Maceió, Alagoas, Brazil

Abstract. Social query is the practice of sharing questions through collaborative environments. In order to receive help, askers usually broadcast their request to the entire community. However, the prerequisite to receive help is to have the problem noticed by someone able and available to answer. Some works found 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 attention of helpers. Our proposal is to analyse CQA history to identify the similar characteristics of previously asked questions that were answered. We believe that adding these characteristics in new questions will impact the receiving of answers. We evaluate our proposal using real world data and a real world experiment. Our results indicate that including “good characteristics” in the question reduce time for first response and improve answer quality.

Keywords: Social query · Community question answering site · Question redesign · Response rate

1 Introduction

The practice of sharing questions through social media is known as social query. Sharing questions on the Web is an ancient way to find information that emulates the Village Paradigm [1]. It originates in forums during early years of internet use [1]. Community Question and Answering sites (CQA) are collaborative environments entirely dedicated to asking and answering questions practice [2].

The most common sharing strategy is broadcasting the problem to everyone in the community. The prerequisite to receive help is that someone able and available notices the questions [3]. However, there is no guarantee if this will happen neither when. Thus, to facilitate this process, researches have been using query routing, i.e., connecting questions and answerers [4]. This could mean recommending questions to potential answerers [5] or recommending experts to questioners directing their requests to [3]. Question routing is an effective way to attract the attention of someone [6].

However, there are questions that are broadcasted that still receive answers. Thus, directing questions is not the only way to find help. Some studies found a correlation between the characteristics of the question and the outcome of receiving or not receiving answers [7–9]. 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 information to improve their chances of finding help.

Thus, in this work, we investigate how adding certain “good” characteristics affects the performance of questions shared through CQA. Our goal with this study is to verify the following claims:

- C_1 – *Questions that receive answers have good characteristics.*
- C_2 – *Questions that do not receive answers do not have good characteristics.*
- C_3 – *Questions with good characteristics will attract more attention than questions with the opposite characteristics.*
- C_4 – *Questions with good characteristics will receive more answers than questions with the opposite characteristics.*
- C_5 – *Questions with good characteristics will receive answers earlier than questions with the opposite characteristics.*
- C_6 – *Questions with good characteristics will receive answers with higher quality than questions with the opposite characteristics.*

To check our claims, we performed two case studies. The first study consisted in sharing the so called “good” questions on real CQA. The second study consisted in analyzing a sample of question regarding the presence and absence of these “good” characteristics. Our findings suggests that questions with good characteristics are answered earlier, receive high quality answers and less requests for clarifications. We aim to use these results to design an interface that computing students can use to share “good” and “attractive” Programming related questions.

The remainder of this paper is organized as follows: Sect. 2 presents Related Work; Sect. 3 describes the context of our study, as well the results of our preliminary investigation through its history; Sect. 4 is about our Evaluation presenting Methodology, Results, Discussion and Threats to Validity; finally, Sect. 5 ends with Conclusion and Future Work.

2 Related Work

The usual social query strategy is broadcasting the question to everyone. However, this is not the best way of taking advantage of the architecture of the environment. After posting a question that will be visible to everyone, there are some struggle scenarios: (1) receiving several responses, (2) receiving wrong or contradictory responses, and (3) to keep receiving responses when no longer needed. Moreover, there is the possibility of receiving no answers because potential responders may never see the question [4].

The prerequisite to receive an answer is someone able and available to help notices the request [4]. The researches about social query usually propose the query routing as the only way to secure help. The process of directing questions to appropriate helpers is known in literature as query routing (this could mean (1) recommending questions to answerers or (2) recommending answerers to the questioner). Thus, directing questions is an attempt to attract someone's attention [3]. Nichols and Kang [6] confirmed that directing questions significantly increases the response rate. However, there are questions that are directed to nobody that still receive answer.

Some studies found a correlation between the characteristics of the question itself and the fact of receiving or not an answer. Burke et al. [10], for instance, found that, in Usenet groups, introductions referencing lurking and a personal connection to the topic of discussion increase the likelihood of getting a reply. In Yahoo! Answers¹, Yang et al. [11] found that medium length questions are less likely to get answered, as well questions from "other" category or with low similarity with their assigned category. According Asaduzzaman et al. [12], the top five reasons to question remain unanswered in Stack Overflow² 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".

Regarding personal social networks' studies, Teevan, Morris and Panovich [9] found that, in 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 [13], 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 received response: questions with extra sentence are less likely to receive "yes/no" answers or requests for clarification. Lampe et al. [14] found that the question type affect the performance of questions shared on Facebook³. Recommendation posts receive more responses than any other question type; while Favor requests usually take a long time until receive a first response. Comarella et al. [15] conducted a study to understanding factors that affect response rate in Twitter⁴ 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.

All these findings could be used to improve the likelihood of one getting answers [5]. 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 [5]. In addition, this could be used to improve questions

¹ <http://answers.yahoo.com>.

² <http://stackoverflow.com>.

³ <http://facebook.com>.

⁴ <http://twitter.com>.

quality and consequently answer quality [16]. The goal of teaching students to ask better questions was explored in [17]. Results revealed a significant difference 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, we could identify common characteristics among answered questions. While users are phrasing new questions, we could suggest to them to 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 community.

3 Investigating CQA History

We used a Brazilian CQA about programming called GUJ⁵ as the context of our study. This is the larger 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 the list of most recent threads. 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. Figure 1 shows an example of a question shared through GUJ.

We split Fig. 1 in areas: (A) the question title; (B) the question tags; (C) the questioner identification; (D) the publishing time; (E) the question description; (F) the social functions buttons (like, share, flag and favorite); (G) the reply

H	I	J	K	L	M	N
criado	última resposta	respostas	visualizações	usuários	link	
2 dias	3 17 horas	3	32	3	1	

Fig. 1. Example of question thread shared on GUJ

⁵ <http://www.guj.com.br>.

button; (H) the questioner identification and publishing time; (I) the publishing time of last answer and answerer identification; (J) the amount of answers received; (K) the amount of views received; (L) the number of users attracted by the question (including questioner); (M) the number of links published through answers; and (N) the list of users that interacted and the amount of contributions of each one.

We conducted a qualitative study in order to identify which are the most common characteristics in answered questions. We started gathering a sample of questions from GUJ containing both answered and unanswered threads. 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 outline a list of characteristics, we conducted 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 help newcomers to ask “good” questions. We confronted and combined both analysis and it resulted in the following list of characteristics which a question can 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 cannot be too short or too long. Regarding with the good title characteristics, users should prioritize: (I) understandable title; (II) medium size title; and (III) a title coherent with the question description subject.
- **Description related characteristics** – After he has been attracted by the title, the potential helper will read the problem description. We observed that some users do not want “waste their time” looking 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 be concise and clear is not always an easy task. Regarding with good description characteristics, users should prioritize: (IV) understandable description; (V) avoid too long description; (VI) showing an example, but avoiding too much code; (VII) avoid description with code only; and, (VIII) when including links, combining them with partial content.
- **Behaviour 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: (IX) use of proper language; (X) including greetings; (XI) avoid be impolite; (XII) avoid be demanding; (XIII) restricting the question to a single problem; (XIV) avoid creating duplicated questions (this can be reached by searching in the community for a similar question, before create a new one); and (XV) avoid create factoid questions (this kind of problem is well solved through search engine use).

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

4 Evaluation

We believe that adding an “Assistance Phase” to help users, before they disclose their problem in a social environment, can improve question quality and response rate. To validate our approach we test the performance of some “enhanced” questions shared on CQAs. In addition, we compared data from a sample of answered and unanswered questions regarding the presence of the good characteristics. Our results indicate that following the suggestions improve response rate, time for the first response, question quality, response quality and question attractiveness.

4.1 Methodology

To test our claims, we planned two studies. The first study consists in the sharing of questions with these good characteristics and with the opposite characteristics and comparing the performance of both groups. The second study consists in the comparison between the data about answered and answered questions on GUJ, regarding the presence and absence of these characteristics.

Our first study works as a concept proof of our belief that adding certain characteristics will impact what happens to the question after being broadcasted. Basically, we formulate 5 questions with the “good characteristics” and 5 questions with the “bad characteristics” (this means do the opposite of the suggestions). Then, we shared these questions during a week in GUJ and analyzed what succeeded: how much people was attracted to the questions, how much people answered the questions, how long take for them to receive answers, the quality of these answers, etc.

All these questions were based on the main problems faced by students during the classes of “Programming I” and “Data Structure and Algorithms”. According [12], too hard questions are more likely to remain unanswered. Thus, while the question topics have a wide range, question difficulties were only low and medium. Regarding answer's quality, although there are many researches in this area [18], since we shared few questions and we would individually define the answers quality, to compare the performance of our questions regarding the answers they received, we used a scale considering all types of answers that we received in ascending order of utility: -2 means that “respondent acted aggressively with the questioner”; -1 means that “respondent did not comprehend the question and asked for more information”; $+1$ means that “respondent only suggested to consult a link”; $+2$ means that “respondent offered a partial/incomplete solution”; $+3$ means that “respondent offer a complete and correct solution”. This scale also considers respondent's effort to provide an answer.

The second study consisted in analyzing a sample of question regarding the presence and absence of these “good” characteristics. Our goal was to investigate if questions answered have the characteristics that we believe attract answers and if questions unanswered miss these characteristics. We sample 100 questions of each type and describe then for the presence and absence of these characteristics.

4.2 Results

Table 1 shows the performance of the questions that we shared for the first study. We also add a question index column to refer these questions later—GQ stands for good question and BQ stands for bad question.

Table 1. Performance summary of questions shared on GUJ

Question Index	Difficulty	Topic	Publishing Time	Time for First Response	Time for the Best Response	People Attracted	Amount of Responses	Responses Quality
GQ1	Low	String API	06:37 PM	20 min	157 min	38	2	[2,3]
GQ2	Low	File API	04:18 PM	13 min	13 min	22	1	[1]
GQ3	Medium	Reflection API	11:48 PM	22 min	22 min	18	1	[2]
GQ4	Medium	Theory	10:49 AM	15 min	350 min	16	2	[-1,3]
GQ5	Medium	Generics	11:09 AM	16 min	16 min	8	1	[-1]
BQ1	Low	Theory	01:47 AM	2032 min	2032 min	47	1	[-2]
BQ2	Medium	XML exception	02:52 PM	-	-	-	-	-
BQ3	Medium	Algorithm	11:43 PM	-	-	-	-	-
BQ4	Medium	Algorithm	06:41 PM	248 min	248 min	25	2	[-1,1]
BQ5	Medium	Swing API	03:34 PM	493 min	493 min	19	1	[2]

These results will be discussed in next section, but we can notice that “bad” questions clearly wait longer for a first response. In addition, answer quality of “good” questions is usually higher.

Figure 2 describes our sample of questions, for the second study, regarding the presence and absence of the “good” characteristics. We used the same Roman numerals of Sect. 3.

We further discuss all these data in next section. But, what we can observe is that real users try to add to their questions good “characteristics”, without even realizing it. In addition, we can only perceive a slightly higher presence of the good characteristics in answered questions group.

4.3 Discussion

Unfortunately, we still do not have data to statically verify our claims; however, observing the results of both studies, we were able to realize interesting patterns that give us directions to qualitatively check them.

Our first claim (C_1) was that “*Questions that receive answers have good characteristics*”. Although, this seems intuitive, data from questions gathered,

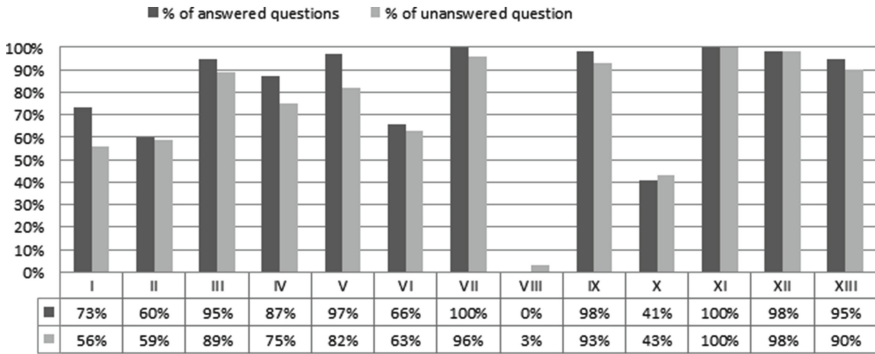


Fig. 2. Data from a sample of answered and unanswered questions

summarized on Table 1, show that both answered and unanswered questions usually have similar characteristics. However, CQAs are environment with high flow of new threads and GUJ is not different. Our first study shows that, when we have good and bad questions mixed, the first type usually receives answer quickly and has less chance of been lost in the thread flow, as it happens with BQ2 and BQ3. Thus, we believe that good characteristics are correlated with the receiving of good answers.

Our second claim (C_2) was that “*Questions that do not receive answers do not have good characteristics*”. This claim is not true, since Fig. 2 shows that even unanswered questions have good characteristics too. However, what we perceived is that, when a question has bad characteristics, it is likely to not receive many answers. In addition, questions that are hard to comprehend have more chances of receiving answers asking for clarifications.

It was not possible to say something about our third claim (C_3) that “*Questions with good characteristics will attract more attention than questions with the opposite characteristics*”. Since GUJ allows us to see how many times people saw the question, from the first study, we were able to perceive that all questions attract almost the same number of users. However, BQ1, one of the first questions published attract a lot of users and it received just a single answer. We believe that this happens due the characteristics of the question that probably annoy most users, including the user who answered. He said “If you break your question and search here or on Google⁶ you will find great material about it. Stop being lazy and search by yourself. When you have a specific question of how to do something you ask here”. The question broke the social contract of the community, when asking factoid questions that have easy answers though search engines, thus, attracting a lot of users, but none useful answer. In addition, questions usually will stop attract people, as times goes by, since they are losing positions on the most recent topics. Unless people keep engage in helping each other in the thread.

⁶ <http://www.google.com>.

Regarding our fourth claim (C_4), “*Questions with good characteristics will receive more answers than questions with the opposite characteristics*”, it was not possible to verify that either. In our first study, most questions received almost the same amount of answers, one or two. Although BQ2 and BQ3 did not receive answers and, probably, this was related to the poor and to long description, respectively, in the second study, answered and unanswered questions had similar characteristics. In addition, we believe that questions will stop receive answers when the community realizes that one already offer a satisfactory solution.

The fifth claim (C_5) that “*Questions with good characteristics will receive answers earlier than questions with the opposite characteristics*” was considered true. Seeing the performance of good and bad questions in Table 1, regarding to the time of the first response, it is clear that the first group receives answers earlier. This, probably, could be explained for the additional effort to answer a “bad question”. The first study showed that poorly written questions usually receives request for clarifications, before receiving a satisfactory answer.

The sixth claim (C_6), that “*Questions with good characteristics will receive answers with higher quality than questions with the opposite characteristics*”, was also considered to be truth. When we have good and bad questions mixed, the first type usually receives answer quickly and has less chance of receiving request for clarification. In addition, as BQ1 shows us, good questions have less chance of receiving negative responses from the community.

4.4 Threats to Validity

This section, briefly, discuss about 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; and the internal validity corresponds to check if the results are a consequence of the manipulation that was done and no others factors. The conclusion validity refers to the correct correlation between what was verified (measured) and the conclusions reached; and, finally, construct validity regards to problems in the design and control of the experiment.

Related to the external validity, a threat to the conclusions reached is that we tested our claims in the context of a CQA about programming. Although, we are confident that our proposal fits any collaborative environment, we cannot guarantee that the same observations will happen outside GUJ.

Related to internal validity, it is true that our observations are highly connected with the questions that we asked, the time when they were released by moderators and who was online on that moment. We tried to use a real world experiment, in order to obtain real world feedback. Sure this brings random factors to the table that could affect the performance of questions shared. However, experiments with historical data from CQA lack these types of “noise”, that are important in real world situations.

Related to conclusion validity, problems may occur on the findings if late answers would come or if we establish the wrong “good characteristics”.

However, we believe that this is very unlikely, since the flow of new threads on G.U.J. is intensive, as time goes by, lesser chances a question thread has to receive more replies; and, regarding to the set of “good characteristics”, it was based in a mixed study involving literature review and classification task.

Related to construct validity, to our first study, we used ten questions asked by students through Java and data structure classes. We enhance these questions with good and bad characteristics. We are aware that maybe most questions shared on CQAs has mixed characteristics and a wide range of topics and complexities, but, as we was trying to analyze the impact of good characteristics only, we could not use too hard questions or mixed characteristics questions.

5 Conclusion and Future Work

In this work, we propose the inclusion of “Assistance Phase” in query routing process. The goal of this new phase is including in the questions characteristics which were present on previously asked answered questions. In addition, it could be used to reduce scope of Expert Search. We used two studies to evaluate the impact of our proposal, while broadcasting a question.

Our findings suggests that questions with good characteristics are answered earlier, receive high quality answers and less requests for clarifications. In addition, we find interesting patterns related to the presence of “opposite” characteristics. This work represents a first step towards the goal of assisting askers in posting attractive questions, by helping them redesign their question, aiming to improve response rate and time for first response.

To future work, we propose analyzing the individual impact of each good characteristic in the receiving of answers, use statistical analysis to check our claims and also test if questions with bad characteristics remain answered for a long time.

Acknowledgements. We want to thank the support of TIPS (<http://www.grouptips.org>) and IFPB–Campus Monteiro regarding the conduction of this study.

References

1. Mansilla, A., Esteva, J.: Survey of social search from the perspectives of the village paradigm and online social networks. *J. Inform. Sci.* **39**(5), 688–707 (2013)
2. Raban, D., Harper, F.: Motivations for answering questions online. *New Media and Innovative Technologies* (2008)
3. De Souza, C.C., De Magalhães, J.J., De Costa, E.B., Fechine, J.M.: 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 (2012)
4. Furlan, B., Nikolic, B., Milutinovic, V.A.: Survey of intelligent question routing systems. In: *Proceedings of the 6th IEEE International Conference Intelligent Systems*, pp. 014–020 (2012)

5. Dror, G., Koren, Y., Maarek, Y., Szpektor, I.: I want to answer, who has a question? Yahoo! answers recommender system. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1109–1117 (2010)
6. Nichols, J., Kang, J.: Asking questions of targeted strangers on social networks. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work, pp. 999–1002 (2012)
7. Dror, G., Maarek, Y., Szpektor, I.: Will my question be answered? predicting “question answerability” in community question-answering sites. In: Proceedings of the Machine Learning and Knowledge Discovery in Databases, pp. 499–514 (2013)
8. Comarella, G., Crovella, M., Almeida, V., Benevenuto, F.: Understanding factors that affect response rates in twitter. In: Proceedings of the ACM SIGWEB Conference on Hypertext and Social Media, pp. 123–132 (2012)
9. Teevan, J., Morris, M., Panovich, K.: Factors affecting response quantity, quality, speed for questions asked via social network status messages. In: Proceedings of the International Conference of Weblogs and Social Media, pp. 630–633 (2011)
10. Burke, M., Joyce, E., Kim, T., Anand, V., Kraut, R.: Introductions requests: Rhetorical strategies that elicit response in online communities. In: Proceedings of the 3rd Communities and Technologies Conference, pp. 21–39 (2007)
11. Yang, L., Bao, S., Lin, Q., Wu, X., Han, D., Su, Z., Yu, Y.: Analyzing predicting not-answered questions in community-based question answering services. In: Proceedings of the 25th AAAI Conference on Artificial Intelligence, pp. 1273–1278 (2011)
12. Asaduzzaman, M., Mashiyat, A.S., Roy, C.K., Schneider, K.A.: Answering questions about unanswered questions of stack overflow. In: Proceedings of the 10th Working Conference on Mining Software Repositories, pp. 97–100 (2013)
13. Teevan, J., Morris, M.R., Panovich, K.: Does anyone know how to get good answers? how social network questions shape replies. Technical Report, No: MSR-TR-2013-62 (2013)
14. Lampe, C., Gray, R., Fiore, A., Ellison, N.: Help is on the way: patterns of responses to resource requests on facebook. In: Proceedings of the 2014 Conference on Computer Supported Cooperative Work, pp. 3–15 (2014)
15. Comarella, G., Crovella, M., Almeida, V., Benevenuto, F.: Understanding factors that affect response rates in twitter. In: Proceedings of the ACM SIGWEB Conference on Hypertext and Social Media (HT 2012), pp. 123–132 (2012)
16. Baltadzhieva, A., Chrupała, G.: Question quality in community question answering forums. In: Proceedings of the ACM SIGKDD Explorations Newsletter, pp. 8–13 (2015)
17. Sullins, J., Mcnamara, D., Acuff, S., Neely, D., Hildebrand, E., Stewart, G., Hu, X.: Are you asking the right questions? The use of animated agents to teach learners to become better question askers. In: Proceedings of The 28th International Flairs Conference, pp. 479–481 (2015)
18. Furtado, A., Andrade, N., Oliveira, N., Brasileiro, F.: Contributor profiles, their dynamics, their importance in five qa sites. In: Proceedings of the 2013 Conference on Computer Supported Cooperative Work, pp. 1237–1252 (2013)