Why Stack Overflow Fails?

Preservation of Sustainability in Community Question Answering

Ivan Srba and Maria Bielikova

Faculty of Informatics and Information Technologies
Slovak University of Technology in Bratislava
Ilkovičova 2, 842 16 Bratislava, Slovakia
ivan.srba@stuba.sk, maria.bielikova@stuba.sk

Abstract — **Enormous amount of knowledge sharing occurs every day in Community Question Answering (CQA) sites, some of which became popular also among software developers and end users (e.g. Stack Overflow or Ask Ubuntu). In spite of their overall success, we can witness emerging problems in some CQA systems – an increasing failure and churn rate. In order to investigate this trend, we conducted a case study focused on Stack Overflow. At first, we evaluated a community perception that indicates that the emerging problems are highly related to the growing amount of low quality content created by undesired groups of users (i.e. help vampires, noobs and reputation collectors). Consequently, we supported these findings by reproducible data analyses of content and community evolution. In order to face the emerging problems, we suggest to provide the users with novel answerer-oriented adaptive support that, in addition, involves a whole community in question answering. These approaches represent an eminent attitude change in the existing question-answering support methods with the aim to preserve a long-term sustainability of CQA ecosystems.**

Keywords— H.3.4.e Question-answering systems; N.3.d Knowledge sharing; H.5.3.c Computer-supported cooperative work

COMMUNITY QUESTION ANSWERING

With increasing popularity of online communities gathered in knowledge sharing systems (e.g. Wikipedia, forums, and mailing lists), their new forms constantly emerge. One of the most popular among them is the concept of Community Question Answering (CQA) sites, such as Yahoo! Answers or Stack Overflow. Members of these communities can ask various questions, which cannot be usually answered easily by standard information retrieval tools [1], while other members can provide answers to them. Besides general CQA systems, various domain-specific communities appeared. Some of them gained high popularity among software developers and end users, such as Stack Overflow, which is even considered as one of the most successful CQA systems ever [2]. Stack Overflow is particularly effective for novices to obtain answers on conceptual or code review questions and, moreover, it can serve even as a supplement for official software documentation if it does not exist or is only very sparse [3].

Currently, the existing CQA systems are perceived mainly as a successful example of collective intelligence due to their high popularity, millions of answered questions, fast question answering process as well as universal availability. In spite of that, some CQA systems become not as successful as they used to be. As an example, we can take the recent evolution of composition of content and community on Stack Overflow.

While the total number of new questions asked each month was growing gradually from the beginning, we witness a noticeable change in 2014 (see Figure 1). At first, we can a see a sharp peak in the total number of new questions around March 2014 followed by a significant drop of 13% over the next three months. At the same time, the number of questions for which information needs of askers were not fulfilled (i.e. questions either deleted due to their poor quality/violation of community rules, or unanswered for more than one month after their posting) overcomes the number of questions for which askers' information needs were fulfilled (i.e. questions with accepted answer - AA). In the second view by means of relative proportions, we can see the constantly increasing proportion of these deleted or unanswered questions among all new questions, which we denote as a *failure rate*. It is growing rapidly – in 2011, three years after site's establishment, it was only 22.45% while in 2014, it was even 39.43%. The development of the failure rate can be precisely predicted by a linear regression with a high significance ($0.192 + 0.0048x$; where x corresponds to the order of month starting from January 2011; $F(1,46) = 3144$, $p < 0.001$, $R^2 = 0.986$). It means that the failure rate increases in average by 0.48% each month.
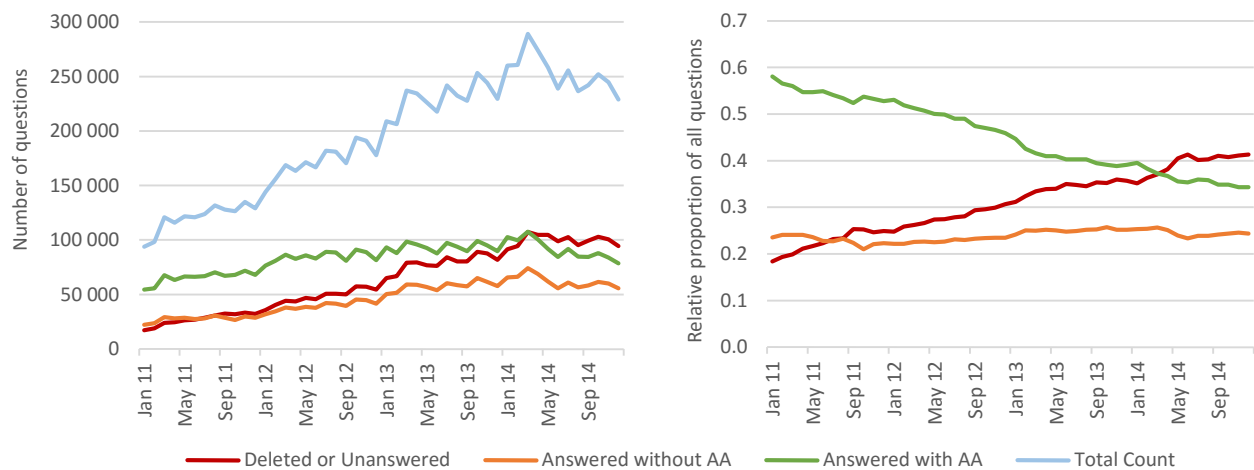

Fig. 1. Evolution of answering success of new questions posted each month [source: Query 1].

Not only content but also community behavior is constantly changing. Only a small fraction of the whole community actively participates in question answering opposing the majority of lurkers (non-contributing members of the community). In 2011, about 15.18% of all registered users posted at least one question or answer each month in average, while in 2014, it was only 5.05% [source: Query 2]. In the detailed view on active users (Figure 2), we can see that the number of active users was increasing until March 2014 and from

that time, it remains rather constant. The proportion of the most important stable users (users who were and remain active) is steadily decreasing (from 41.05% in 2011 to 34.89% in 2014) in contrast to the proportion of one-time users (users who became active just in one month), which is on the contrary increasing (from 30.80% in 2011 to 33.12% in 2014). As the result, the number of one-time users even overcame the number of stable users for the first time in November 2014. The outflow of stable users is related to a growing number of churn users (users who were active at least once during previous three months and became inactive for following three months; the definition was adopted from Pudipeddi et al. [4]). From April to June 2014, the number of churn users for the first time notably overcame the number of newcomers (i.e. users who became active in the particular month, opposite to churn users). This negative trend is reflected in a *churn rate (*a relative proportion of users who churned in a particular month from all active users), which increased from 12.52% in 2011 to 15.85% in 2014. Similarly as the failure rate, also the churn rate can be modelled by a linear regression ($0.121 + 0.0009x$; where x corresponds to the order of month starting from January 2011; $F(1,46) = 149$, $p < 0.001$, $R^2 = 0.764$). It means that the churn rate increases in average by 0.09% each month.
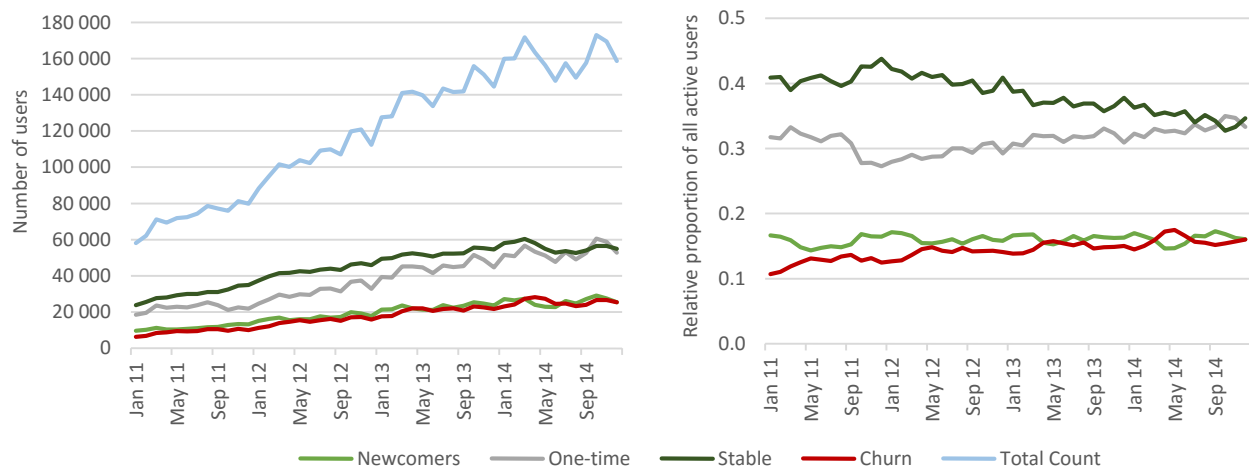

Fig. 2. Evolution of composition of active users [source: Query 3].

Moreover, besides the aggregate numbers also direct feedback from community members in various Internet discussions and blogs (e.g. http://michael.richter.name/blogs/why-i-no-longer-contribute-to-stackoverflow) points out the emerging problems that prevent the long-term sustainability of Stack Overflow. Despite seriousness of this phenomenon (i.e. the increasing failure and churn rate), it has not been well-described yet and furthermore, we are not aware of any particular works aiming to effectively face it. Therefore, we conducted a case study on Stack Overflow to provide a deeper insight into the evolution of CQA communities. Following the obtained results, analyses of state-of-the-art approaches as well as our experiences with an educational CQA system Askalot [5], we propose a shift in providing adaptive collaboration support that can contribute to preserve a long-term sustainability of CQA ecosystems.

CHARACTERIZING CONTENT AND USERS IN CQA

Openness of CQA systems is closely connected with the diversity of users' expertise and activity levels as well as with quality of the created content. This diversity is fruitful for efficient knowledge sharing among people with different levels of expertise, but at the same time, it also prevents CQA systems from becoming trustful archives with entirely unique and quality content. In general, users in CQA systems can be categorized in three dimensions: (1) according to their preferred activities (askers vs. answerers); (2) according to an amount of activity carried out in a system (active vs. passive persons, so called lurkers); (3) and according to their knowledge (level of expertise). Since these dimensions are perpendicular, it is possible to combine them and thus categorize users according to behavior to various user stereotypes. In order to achieve successful question answering, it is essential that the community comprises particular types of users (e.g. active answerers with a high level of expertise). On the other hand, some types of users are not very desirable, although they represent a natural element of each community (e.g. users who ask many low-quality questions). A previous study [6] examined dynamics of these stereotypes on the basis of data from Super User system collected before the end of August 2011. They recognized that the composition of the community is constantly changing – the proportion of some stereotypes in the total number of users is stable while the proportion of other stereotypes (e.g. expert answerers and low-activity users) dynamically changes.

We suppose that the increasing failure and churn rate on Stack Overflow can be explained by a constant evolution of proportionality of content quality and user stereotypes in time. More specifically, we hypothesize that the undesirable types of users and their content has become too widespread and overloaded potential answerers. In spite of several studies, which investigate evolution of individuals' behavior (e.g. churn prediction [4]), the existing research has not focused on time evolution of whole CQA communities and their composition so far. We believe, that this perspective on CQA systems can provide us with deeper insight into the emerging problems.

CASE STUDY: EVOLUTION OF COMMUNITY ON STACK OVERFLOW

In order to verify our hypothesis, we performed a case study on Stack Overflow, where the emerging problems are probably most eminent. In addition, Stack Overflow plays a model role for Stack Exchange platform, which unites a dozen of CQA sites with various topics (e.g. Ask Ubuntu, Mathematics).

Context of the Study. At first, we analyze discussions on Meta Stack Overflow, which is a specific part of Stack Overflow with questions related to the system itself. We manually evaluated 150 questions with the highest number of votes provided by the community (with voting score at least 7). We identified 12 questions, which directly pointed to the negative development of the community. All were posted after March 2014, i.e. after the negative changes in the total number of questions and active users appeared. The importance of these problems is even more highlighted by the fact that 1st and 3rd highest ranking questions (by number of votes as of May 2015) were concerned with these topics: *Why is Stack Overflow so negative*

of late? (http://meta.stackoverflow.com/questions/251758/) and *Question quality is dropping on Stack Overflow* (http://meta.stackoverflow.com/questions/252506/).

Community perception indicates that the decline on Stack Overflow is a result of its increasing popularity and openness, which has resulted into a massive arrival of users with low level of expertise. Consequently the system was flooded with too many questions that were not interesting for other users. Furthermore, the community identified and denominated three groups of users who are the major origin of this undesired content:

1. *Help Vampires* – users who ask a great number of questions without spending any effort to find a required knowledge (e.g. from search engines or archives of already solved questions). Consequently, the posted questions are often tedious or even duplicated. Help vampires are interested only in getting answers to their questions and they do not return the help received back to the community.
2. *Noobs* – users with low level of expertise who create mainly trivial questions with poor quality. Noobs overload the system with a significant amount of low-quality content and make finding of the unique and interesting questions very difficult.
3. *Reputation Collectors* – users who answer as many questions as possible (commonly regardless their insufficient knowledge in the question's topic), primarily to gain the reputation. On one side, these users contribute to the system (by means of assisting experts in answering uninteresting questions), on the other side, they mutually reinforce and motivate help vampires and noobs in asking more low-quality questions.

Opposing these undesired groups of users, another part of Stack Overflow community exists: the *Care Takers* – experts, who want to keep the system clean and with valuable content. Care takers regularly search for interesting questions and provide good answers to them. Their presence is essential and it is important to provide adequate motivation to keep these users active and devoted to the community.

The community–based perception points out that the proportionality of different types of users forming the community has put the CQA ecosystem off balance what also supports our hypothesis. We consider community feedback as relevant, yet we decided to support it by quantitative analyses.

Quantitative Study. We employed a dataset from Stack Overflow which contains all non-anonymous activities. The dataset is publicly available and distributed under creative commons license by means of a dataset dump or by Data Explorer tool, which allows us to investigate data by means of SQL queries. We decided to employ Data Explorer tool to make the analyses presented in this paper easily reproducible at any time (all presented results are accompanied by references to queries that are publicly available and executable at the latest versions of datasets at http://data.stackexchange.com/users/16409/ivan-srba/). Due to this solution, it is possible to use the same methodology and continue in monitoring the further evolution of Stack Overflow as well as all CQA systems in Stack Exchange platform.

At first, we investigated evolution of content quality in time. We consider votes provided by the community as a relatively precise estimation of content quality. However, overall score (a difference between positive and negative votes) can be influenced by the various time lapse between content creation and time when the analyses are performed. Thus to preserve reproducibility of our analyses, we take into consideration only those votes that were created within one month after the content was posted. In addition, number of votes can be influenced by popularity of particular topics. Therefore, to estimate quality of questions/answers more accurately, we normalized their score by a mean score obtained by other questions/answers with the same user-assigned tags (if a post has more than one tag assigned, the score is normalized by means of each of these tags separately and an average from all partially normalized scores is calculated). According to this score, we divided questions and answers into four groups: low quality (with negative score); neutral quality (with score equal to 0); good quality (with positive score which is below 1.5 multiple of the average score); and high quality (with positive score above 1.5).

The absolute numbers of new questions asked each month with good and high quality remained stable regardless the long-term increase in the total number of all questions (see Figure 3). As the result, their relative proportions decreased from 30.79% in 2011 to 18.43% in 2014 and from 5.96% to 2.02%, respectively. This decline is associated with the growth in the number as well in the proportion of low-quality questions that constituted only 4.99% of all questions in 2011 and in 2014, it was even 16.72% (the greatest rise was between April and May 2014 by 5.24% when also the number of low-quality questions overcame the number of good-quality questions). The proportion of neutral questions was relatively stable during all four years (59.94% in average). It means that the most of the increase in the total number of questions is created by uninteresting questions with zero or even negative score. This finding confirms our hypothesis as well as the community perception that the system is flooded by content that nobody cares about, while really interesting content is getting rarer.
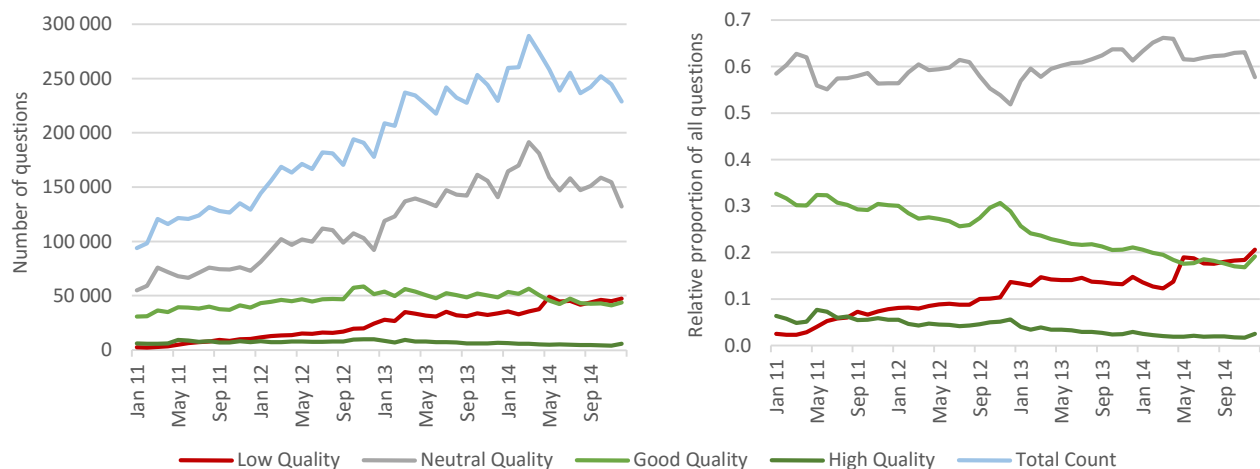

Fig. 3. Evolution of questions' quality [source: Query 4].

The increasing number of questions is not, however, reflected in the number of answers (see Figure 4). Quite the opposite, it is possible to notice a three-month long decrease by 23.90% in the number of answers from April to June 2014 (there are also other declines, e.g. between May and June 2013 by 8.73%, nevertheless they are not so extensive). During this rapid decrease, the answering capacity of Stack Overflow community returned to an average achieved in 2012 (about 280 thousands of answers per month). This indicates that the community was not able to handle so many incoming (mainly low-quality and uninteresting) questions what finally resulted in the continuous increase of the failure rate. Relative proportions reveal the decrease of good and high-quality answers as well as the increase of low-quality answers, however, they are not as significant as for questions.
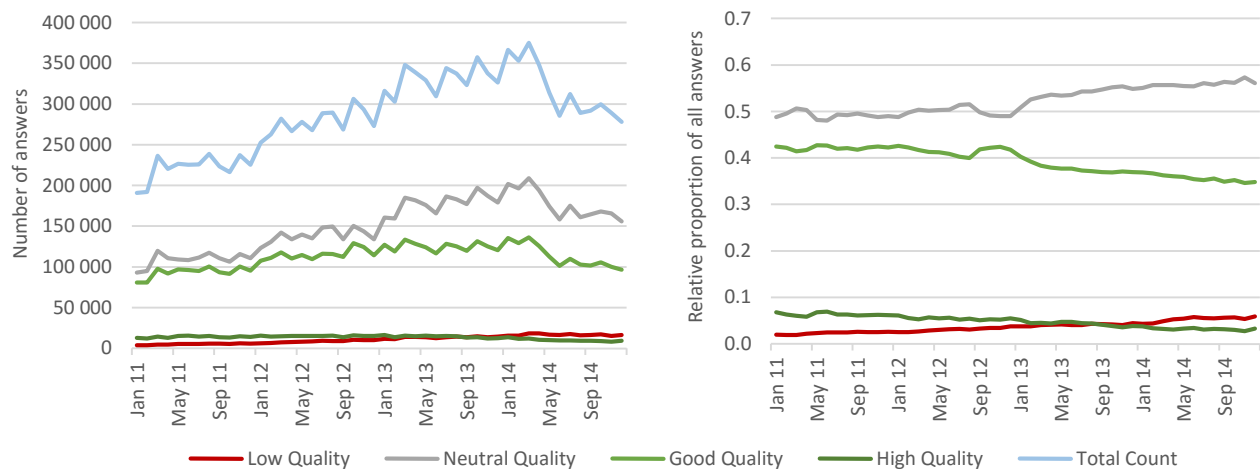

Fig. 4. Evolution of answers' quality [source: Query 5].

Secondly, we investigated the evolution of quantity of user stereotypes, which were denoted by the community. To assign stereotypes to users, we employed a set of rules that are based on the previously provided characteristics, which we additionally supported with the results from Furtado et al. [6]: 1) *help vampire* is a user who asked during a particular month at least two questions and did not provide any answer; 2) *noob* is a user who asked at least of two questions of low or neutral quality; 3) *reputation collector* is a user who provided at least three answers on low-quality questions and was systematically active during at least three days; and finally, 4) *care taker* is a user who was similarly active at least three days and provided high-quality answers on high-quality questions. We identified a long-term outflow of care takers (see Figure 5). Despite the overall community growth, the proportion of care takers among all active users decreased from 3.70% in 2011 to 1.24% in 2014. The proportion of help vampires is relatively stable with average about 13.15%. On the other side, the proportion of noobs increased from 6.23% in 2011 to 10.11% in 2014. Reputation collectors also became more common from 4.11% in 2011 to 5.98% in 2014. These results support the community perception that important care takers tend to leave the system. Also we confirmed

the rise of noobs and reputation collectors, nevertheless, it seems that help vampires always were a natural element of CQA community.
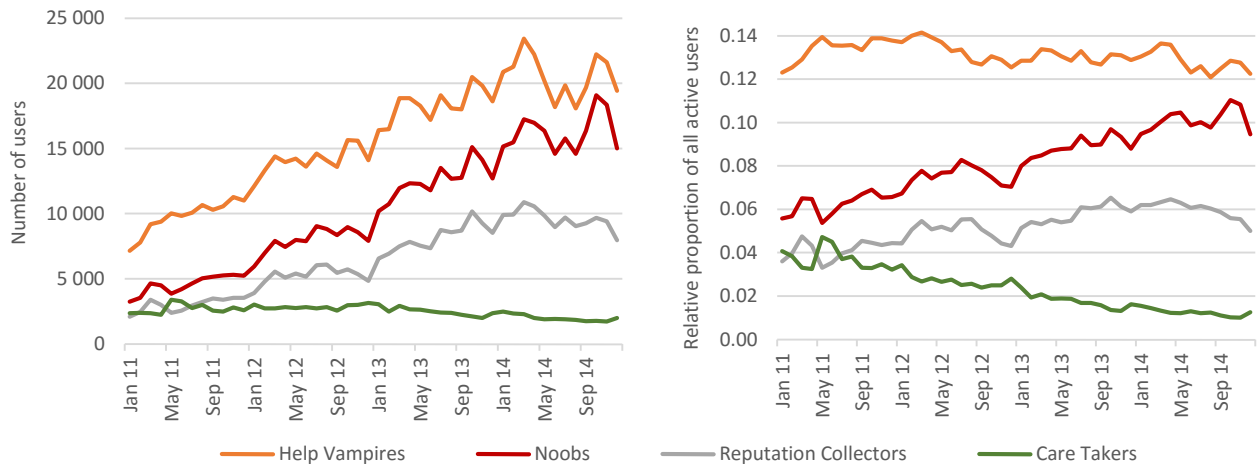

Fig. 5. Evolution of stereotypes assigned to active users [source: Query 6].

In order to evaluate the relation between the failure/churn rate and content/user evolution numerically, we calculated cross-correlation between time series representing the failure/churn rate (see Figure 1 and 2) and relative proportions of each kind of content/users (see Figure 3, 4 and 5). To calculate cross-correlation, the R script was employed (cross correlation was calculated by means of *ccf* function available in *pracma* library; lag was set to -1 month to measure how content/user evolution leads the failure/churn rate; in addition, a linear trend was removed from each time series by *detrend* function). It means that the obtained results measure how well content/user development can predict changes in the failure/churn rate in the following month.

The significant cross-correlations confirm that low-quality questions and low-quality answers lead to the higher failure rate (see Table 1). Moreover, the correlations obtained for the churn rate confirmed that the proportion of help vampires, noobs and reputation collectors leads to higher user churn. On the other side, good and high quality content keeps users active and devoted to the community, nevertheless, the correlations are in these cases just a little bit below the significance level.

Table 1. Cross-correlations between time series representing the failure/churn rate and proportion of various kinds of content/users analyzed in the quantitative study (significant correlations at 5% level of significance are highlighted in bold).

		Failure Rate	Churn Rate
Questions	Low Quality	**0.422**	-0.112
	Neutral Quality	-0.148	0.243
	Good Quality	-0.149	-0.254
	High Quality	0.041	-0.215
Answers	Low Quality	**0.444**	0.024
	Neutral Quality	-0.035	0.121
	Good Quality	-0.093	-0.065
	High Quality	0.078	-0.270
Users	Help Vampires	0.071	**0.558**
	Noobs	0.122	**0.423**
	Reputation Collectors	-0.027	**0.325**
	Care Takers	0.034	-0.239

To conclude, the results from the quantitative study confirmed the community perception that the low-quality content and the undesired user groups are closely connected with the emerging problems. Thanks to the common database structure in the whole Stack Exchange platform, Data Explorer tool allowed us to perform the same analyses also with other CQA systems. For those systems that are recently undergoing through a rapid expansion similarly as Stack Overflow previously did (e.g. Ask Ubuntu), we identified very similar trends.

A SUGGESTION TO PRESERVE CQA SUSTAINABILITY

We identified several possibilities how to deal with the emerging problems. One option is to change community rules and restrict the overall openness of CQA systems (e.g. limit a number of questions which users can ask per week). However, this solution will solve the problems only partially (as there would be still low-quality content, although its amount would be reduced) and temporally (as restrictions often pose new unexpected problems). We emphasize that it is not possible to get completely rid of low quality content. Instead of that, we suggest that CQA systems should take various content quality and users' expertise into consideration.

Only very recently, approaches to automatic detection of low quality posts [7] and content abusers [8], [9] have been proposed to help moderators with deletion of these posts or banning these users, respectively. These solutions might be really effective, however, in spite of their overall precision (varying from 0.70 to 0.80), they can misclassify good quality content or innocent users and as the result, taking the extreme actions can finally lead to even more undesired and antisocial behavior [9]. Therefore, we suggest to solve the

emerging problems in alternative and less invasive ways. At first, it is possible to adjust a reputation system to reflect a value of contributions more accurately and thus motivate users to provide good answers on good questions (e.g. reputation received for providing an answer can reflect corresponding question difficulty). We introduced an example of this kind of reputation system in our recent paper [10].

Another possibility, which we consider even more promising, is to provide users with an appropriate adaptive collaboration support. In the CQA domain, several approaches have been already proposed to adaptively support effective knowledge sharing (e.g. personalized recommendation of questions [11]). Unfortunately, adaptive support methods still do not reflect the emerging problems appropriately. Moreover, some of these methods even indirectly support the undesired user groups by giving preference to their goals instead of other users (e.g. care takers). For preserving a long-term sustainability of CQA communities, it is, however, necessary to change an attitude in providing an adaptive support. We propose two basic approaches how to achieve this shift.

Answerer-oriented Approaches – At first, a majority of the existing adaptive support methods can be characterized as asker-oriented as they are either explicitly dedicated to askers or they are primarily focused on askers' goals while answerers' preferences and expectations are suppressed.

The asker-oriented approach is visible especially in question routing methods (i.e. recommendation of new questions to potential answerers). Most of the existing methods recommend questions to experts regardless of the real question difficulty (and required minimum level of expertise for proper answering). This approach is really successful in achieving askers' goals (to receive a high quality answer), although it completely overlooks those experts who prefer to answer more difficult and challenging questions within their limited time capacities. In addition, existing methods usually route questions to potential answerers with the same or very similar topics and thus answerers can lose motivation easily. In order to prevent this filter bubble and to meet answerers' expectations, diversification of recommendation should be applied [11].

Involvement of the Whole Community – Another drawback of existing adaptive methods is that they involve and motivate only a small portion of community to actively participate in the question answering process. To maintain CQA ecosystem, it is necessary to satisfy expectations of all types of users [11].

For illustration, we continue with an example of the question routing problem. As the existing methods prefer users with a high level of expertise, other users are involved only very rarely while experts are easily overloaded and capacity of other users is left unutilized. Trivial questions (especially those asked by help vampires or noobs) can be usually answered by other users who have sufficient knowledge, but are not necessary experts with the highest level of expertise. Moreover, existing methods usually rely on previous user activities in the CQA system and thus they are not able to route questions to newcomers (due to a well-known cold start problem) or to lurkers (due to lack of their sufficient activity).

In our previous work [12], we demonstrated a complex example how the given recommendations to preserve CQA sustainability can be applied in the case of question routing. We introduced a novel question

routing method which involves all possible answerers by employing their public non-QA data (i.e. about me descriptions, blogs) in order to supplement QA data (i.e. data from the question answering process in CQA system itself) and thus provide question recommendations more precisely and even to all kinds of users.

To sum it up, positive outcomes of CQA systems (e.g. the number of questions, the average time to the first answer and the great archives of answered questions) outperform the emerging problems that are reflected in the increasing failure and churn rate. Their negative impact is, however, significantly growing in the recent time. A set of easily executable analyses, which we published in the Data Explorer tool, allow to monitor the future development of Stack Overflow and other CQA systems based on Stack Exchange platform. The conclusion from the performed study is that openness of CQA communities leads to increasing numbers of users who create mainly low-quality and uninteresting questions. As the result, these questions remain very often unanswered and demotivate experts who slowly leave the community.

We propose the attitude change in adaptive support methods to deal with the emerging problems and thus to contribute to the long-time sustainability of CQA ecosystems. This shift aims to prevent expert overloading, to make answerers more satisfied and to optimally utilize knowledge embedded in all community members, and can be described by two approaches:

1. instead of giving a focus only on askers and their goals, preferences and expectations of answerers should be considered as well;
2. instead of involving only a subset of active and expert users, a whole community should be engaged in a question answering process.

REFERENCES

[1] Q. Liu, E. Agichtein, G. Dror, Y. Maarek, and I. Szpektor, "When Web Search Fails, Searchers Become Askers: Understanding the Transition," in *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '12*, 2012, pp. 801–810.

[2] L. Mamykina, B. Manoim, M. Mittal, G. Hripcsak, and B. Hartmann, "Design Lessons from the Fastest Q&A Site in the West," in *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, 2011, pp. 2857–2866.

[3] C. Treude, O. Barzilay, and M.-A. Storey, "How Do Programmers Ask and Answer Questions on the Web? (NIER Track)," in *Proceedings of the 33rd international conference on Software engineering - ICSE '11*, 2011, p. 804.

[4]   J. Pudipeddi, L. Akoglu, and H. Tong, "User Churn in Focused Question Answering Sites: Characterizations and Prediction," in *Proceedings of the 23st international conference companion on World Wide Web - WWW '14 Companion*, 2014, pp. 469–474.

[5]   I. Srba and M. Bielikova, "Askalot: Community Question Answering as a Means for Knowledge Sharing in an Educational Organization," in *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing - CSCW'15 Companion*, 2015, pp. 179–182.

[6]   A. Furtado, N. Andrade, N. Oliveira, and F. Brasileiro, "Contributor Profiles, their Dynamics, and their Importance in Five Q&A Sites," in *Proceedings of the 2013 conference on Computer supported cooperative work - CSCW '13*, 2013, pp. 1237–1252.

[7]   L. Ponzanelli, A. Mocci, A. Bacchelli, M. Lanza, and D. Fullerton, "Improving Low Quality Stack Overflow Post Detection," in *Proceedings of IEEE International Conference on Software Maintenance and Evolution - ICSME '14*, 2014, pp. 541–544.

[8]   I. Kayes, N. Kourtellis, D. Quercia, A. Iamnitchi, and F. Bonchi, "The Social World of Content Abusers in Community Question Answering," in *Proceedings of the 24th International Conference on World Wide Web - WWW '15*, 2015, pp. 570–580.

[9]   J. Cheng, C. Danescu-niculescu-mizil, and J. Leskovec, "Antisocial Behavior in Online Discussion Communities," in *Proceedings of AAAI International Conference on Weblogs and Social Media - ICWSM '15*, 2015.

[10]   A. Huna, I. Srba, and M. Bielikova, "Exploiting Content Quality and Question Difficulty in CQA Reputation Systems," in *Proceedings of International Conference on Nework Science - NetSciX '16*, 2016, to appear.

[11]   I. Szpektor, Y. Maarek, and D. Pelleg, "When Relevance is not Enough: Promoting Diversity and Freshness in Personalized Question Recommendation," in *Proceedings of the 22nd international conference on World Wide Web - WWW '13*, 2013, pp. 1249–1259.

[12]   I. Srba, M. Grznar, and M. Bielikova, "Utilizing Non-QA Data to Improve Questions Routing for Users with Low QA Activity in CQA," in *Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining - ASONAM '15*, 2015, pp. 129–136.

Ivan Srba is currently a doctoral student at the Institute of Informatics and Software Engineering, Slovak University of Technology in Bratislava. His research interests are in the area of web-based systems, which utilize concepts of collaboration and collective intelligence. Contact him at ivan.srba@stuba.sk.

Maria Bielikova is a full professor at the Institute of Informatics and Software Engineering, Slovak University of Technology in Bratislava. Her research interests are in the area of web-based systems with special emphasis on personalization, context awareness and collaboration. She is a senior member of IEEE, a member of its Computer Society, senior member of ACM and member of International Society for Web Engineering. Contact her at maria.bielikova@stuba.sk.

SUMMARY QUESTIONS AND HIGHLIGHTS

1. Why Stack Overflow fails? Answering questions about sustainability in Community Question Answering.
2. Analyses of recent evolution of Stack Overflow revealed new emerging problems: increasing failure and churn rate.
3. Help vampires, noobs, reputation collectors and their influence on successfulness of Stack Overflow.
4. Proposing suggestions to preserve a long-term sustainability of Community Question Answering.
5. Answerer-oriented approaches and involvement of the whole community - how to maintain Community Question Answering.