2021 IEEE/ACM 13th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE) | 978-1-6654-1409-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/CHASE52884.2021.00027

# A Virtual Mentor to Support Question-Writing on Stack Overflow

Nicole Novielli, Fabio Calefato, Federico De Laurentiis, Luigi Minervini and Filippo Lanubile
*University of Bari, Italy*
{nicole.novielli,fabio.calefato,filippo.lanubile}@uniba.it, {f.delaurentiis,l.minervini7}@studenti.uniba.it

***Abstract*—Recent studies on Stack Overflow show that question-related factors, such as conveyed sentiment and presentation quality, can significantly influence the probability of obtaining a useful answer. At the same time, the mentorship provided by human experts was proven effective in support novice users in effective question-writing. In line with previous empirical findings, we developed *QAvMentor*, a tool capable of providing online, real-time, automated mentorship during question-writing in Stack Overflow.**

***Index Terms*—Knowledge sharing, Question Answering, Virtual Mentoring, Stack Overflow**

## I. Effective Question-writing on Stack Overflow

The popularity of community-based question-answering (QA) has deeply affected the way people collectively build knowledge, as demonstrated by the success of platforms such as Stack Exchange, of which Stack Overflow is the most popular site. The effectiveness of QA depends on the will of the community members to provide support by answering questions on the platform. Calefato et al. [1] investigated how developers can increase the chance of eliciting a successful answer to their technical questions on Stack Overflow and provided data-driven guidelines for effective question-writing. On the other hand, Ford et al. [2] showed that it is possible to design and implement *ad hoc* support to novice users. Inspired by the findings of these two studies, which we summarize in the following, we implemented `QAvMentor`,[1] a virtual mentoring tool to support novice users in effective question-writing in Stack Overflow.

### A. Data-driven guidelines for successful question writing

`QAvMentor` implements a set of checks derived from data-driven guidelines on how to write a good question on Stack Overflow proposed by Calefato et al. [1]. They developed a conceptual framework of factors affecting the success of a question in Stack Overflow- The framework distinguishes between actionable factors — namely, Affect, Presentation Quality and Time — on which information seekers can act upon while writing a question, and non-actionable ones, — i.e., Reputation. Affect is measured as sentiment polarity conveyed by a question, while the Presentation Quality is operationalized using a set of quality indicators such as conciseness of the title, appropriate use of capital letters, presence of a code snippet, and so on. Time is included based on empirical evidence from previous research suggesting that low vs. high efficiency time slices exists in the community. The reputation is included as a control factor.

They used logistic regression to quantitatively analyze 87K Stack Overflow questions, in order to assess the impact of actionable factors on the success of questions, i.e. on their probability to obtain an accepted answer. The results of logistic regression show that, they report that, regardless of the user reputation, successful questions are short, contain code snippets, and do not abuse with uppercase characters. As for sentiment, successful questions adopt a neutral emotional style while both positive and negative sentiment correlates to a lower probability of success. Calefato et al. use their findings to derive a set of evidence-based guidelines which we use to guide the design and implementation of our tool.

### B. Mentoring Novice Users in Online Question-Answering

The results of the study performed by Calefato et al. [1] suggest how question-answering novice users are often unaware of community norms. Providing support to novice users becomes crucial also to enable effective knowledge-sharing and building, thus resulting in the creation of long-lasting value QA threads in technical sites. Ford et al. [2] conducted an empirical study to assess the effectiveness of a just-in-time mentorship program to Stack Overflow novice users. During a month-long experiment, 271 participants received on-site support from human mentors, with whom they interacted online in a dedicated Help Room. Stack Overflow mentors, specifically recruited for the purpose of the study, provided suggestions on how to improve questions based on their experience as senior community members. Suggestions addressed a broad range of question quality aspects, from providing context to code formatting, to compliance with code of conduct. They found that the mentoring program proved effective for novice users, with mentored questions scoring on average 50% higher than non-mentored questions. As a further effect of mentoring, they observed a lower rate of off-topic and poor-quality questions.

## II. The QAvMentor tool

`QAvMentor` is a browser extension for Chromium-based browsers [2], which supports novice users in effective question-writing on Stack Overflow. `QAvMentor` verifies that the real-time actionable guidelines are taken into account as users

[1]Backend: https://github.com/collab-uniba/qavmentor, Frontend: https://github.com/collab-uniba/qavmentor-plugin

[2]Available at: https://tinyurl.com/qavmentor

2574-1837/21/$31.00 ©2021 IEEE
DOI 10.1109/CHASE52884.2021.00027

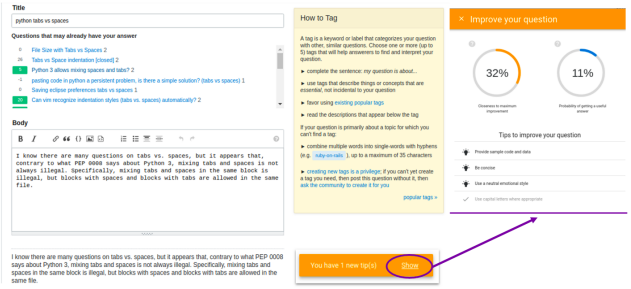


Fig. 1. QAvMentor checks compliance with guidelines during question-writing and informs users that new tips are available (left). Tips are provided (right) with the success probability for the question in its current state.

enter the question title and body. During question-writing, the extension checks if any of the guidelines is not implemented, thus prompting a message indicating the availability of suggestions for improving the question title and/or body. Upon clicking on the tip box in orange (see Figure 1), users receive an estimate of the likelihood for the question in the current form of obtaining a useful answer (the blue circle in figure) as well as a percentage estimate of how close it is to the maximum improvement that could be achieved by following the tips (the orange circle in figure). These tips are listed at the bottom (in black in figure) along with the guidelines that are already satisfied (listed in grey, next to a checkmark). At the time of writing, QAvMentor extracts the metrics associated to the affect, presentation quality, and time dimensions in the framework originally developed by Calefato et al. [1]. Then, it computes the success probability for the question by leveraging the logistic regression model they built in their original study and generates the tips based on the collected metrics. For example, QAvM recommends to *'provide sample code and data'* in absence of a code snippet. Similarly, it suggests to *'use a neutral emotional style'* in presence of lexical cues for either positive or negative sentiment.

## III. Evaluation

We ran a preliminary user study to assess the usability of `QAvMentor` as well as the clarity of the tips provided. We recruited six undergraduate Computer Science students, who individually performed five different tasks. The asks either consisted in question-writing starting from a a code snippet generating an error message, or in question editing starting form draft question that students were requested to change following the tips for improvement.

Upon completion of all tasks, the students reported the ease of use and satisfaction in a debriefing session. They were also invited to provide further comments and suggestions for improvement. Overall, the system was positively evaluated. Some of the participants recommended including suggestions for changes along with tips. Other than analyzing the user subjective feedback, we used further assess the ease of use by computing: (i) the *completion rate*, that is the proportion of tasks completed; (ii) the average *completion time* for each task. All participants successfully completed the five tasks. As for the completion time, the first task required substantially longer to be completed, i.e. 9 minutes on average vs. 1 minute required for the others, probably due to the need for familiarizing with the tool interface during the very first task of the session.

## IV. Discussion and Conclusions

We implemented and deployed `QAvMentor`, a browser extension providing virtual mentoring to novice users in Stack Overflow. `QAvMentor` leverages empirically-driven guidelines [1] to assess question quality and provide *ad hoc* tips for question improvement at the time of writing, that is before the question is posted to Stack Overflow. By publicly distributing `QAvMentor` we aim at encouraging researchers to design and conduct large-scale, empirical studies to further assess and validate empirically-driven guidelines for effective question-writing on Stack Overflow. Furthermore, by releasing the tool as an open-source project, we hope to attract contributions from both researchers and practitioners. Specifically, we envision the integration of finer-grained sentiment classifiers for detecting toxicity [3], anger towards peers [4] and code of conduct violation [5]. The integration of such fine-grained sentiment analysis tools would enable the correct classification of the writing style beyond positive or negative polarity. For example, by correctly distinguishing between hostile behavior and other negative emotions, such as *fear* of failure or *disappointment* for not being able to solve a problem, `QAvMentor` could deliver more effective tips. In the former case, it could guide the users towards the adoption of constructive patterns of interaction and neutral writing style, in line with community standards. As a further extension, a module for duplicate question detection [6] could also be integrated to help novice users browse previously answered questions addressing analogous problems, thus avoiding redundancy. Finally, we hope to foster the engagement and participation of novice users through the large-scale adoption of our virtual mentoring tool.

## References

[1] F. Calefato, F. Lanubile, and N. Novielli, "How to ask for technical help? evidence-based guidelines for writing questions on stack overflow," *Inf. Softw. Technol.*, vol. 94, no. C, p. 186–207, Feb. 2018.

[2] D. Ford, K. Lustig, J. Banks, and C. Parnin, ""we don't do that here": How collaborative editing with mentors improves engagement in social q&a communities," in *Proc. of the 2018 CHI Conf. on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: ACM, 2018, p. 1–12.

[3] N. Raman, M. Cao, Y. Tsvetkov, C. Kästner, and B. Vasilescu, "Stress and burnout in open source: Toward finding, understanding, and mitigating unhealthy interactions," in *Proc. of the ACM/IEEE 42nd Int'l Conf. on Software Engineering: New Ideas and Emerging Results*, ser. ICSE-NIER '20. New York, NY, USA: ACM, 2020, p. 57–60.

[4] D. Gachechiladze, F. Lanubile, N. Novielli, and A. Serebrenik, "Anger and its direction in collaborative software development," in *Proce. of the 39th International Conference on Software Engineering: New Ideas and Emerging Results Track*, ser. ICSE-NIER '17. IEEE, 2017, p. 11–14.

[5] P. Tourani, B. Adams, and A. Serebrenik, "Code of conduct in open source projects," in *2017 IEEE 24th Int'l Conf. on Software Analysis, Evolution and Reengineering (SANER)*, 2017, pp. 24–33.

[6] Y. Zhang, D. Lo, X. Xia, and J. Sun, "Multi-factor duplicate question detection in stack overflow," *Journal of Computer Science and Technology*, vol. 30, pp. 981–997, 09 2015.