Gender Gap on GitHub

E. Jobst

February 9, 2020

Abstract

Female participation, which is already bad in the computer science field, is even worse regarding contributions to FLOSS projects. In this work the current state-of-the-art on the gender gap on the example of GitHub is reviewed. The barriers women face on contributing, range from the consequences of descriptive and prescriptive stereotypes, to women acting differently on GitHub in order to avoid backlash. One major issue is how competent women often lack the self-esteem to contribute. Finally, possible solutions to these problems are presented, with a considerable focus on supporting newcomer.

1 Introduction

Female participation in the software world is already sparse. Using Austria as an example only 26,7% of the female university graduates in the year 2016/17 have chosen a technical study. This is already a slight improvement to the statistics that were conducted 10 years prior, where 20,5% graduated in a technical field [1]. Nevertheless, these numbers decline at a rapid rate regarding contributors of FLOSS projects [12].

In the version-control platform GitHub only 5% of the projects are owned by women. Of the 5000 top-most contributor only around 3% are female [12]. This may seem frustrating as diverse team-structures are very beneficial for a project. Ortu et al. proved that with a gender-diverse team the amount of time spent on solving an issue on GitHub drops significantly [6][11].

With these numbers present it makes someone wonder on the reasons why such a small number of women participate in the FLOSS community with such an, assumingly, small amount of time. The purpose of this paper is collecting issues and barriers that women face on the example of GitHub. The found issues will be contrasted with possible solutions to them.

2 Gender Gap on GitHub

2.1 Contribution Differences of Women and Men

Robles et al. highlight the differences in contribution type between women and men. In their study they analysed it with splitting the types into three categories called „code, programming", „other" and „both". The category „other" is defined with tasks, that are necessary, however, outside the realm of adding new features or fixing bugs, like documentation, testing or translating. It's noteworthy that men mostly contributed via source-code contributions, however, women performed the other tasks more frequently [7].

2.2 Pull-Request Acceptance Rate

Terrell et al. made an in-depth investigation on the pull-request acceptance rate of women. In general, women have a higher pull-request acceptance rate than men. However, they found possible bias in certain circumstances: When the gender is unidentifiable their acceptance rate is higher by 12%, than in comparison to when their gender is known.The rate drops for men as well, however, only by 3,8%. This was compared by splitting the data up in "Insider", which are the owner or collaborator of the project and "Outsider", which are regular contributor. This decline in the pull-request acceptance rate has only been observed with contributor, meaning for owner or collaborator such a decline wasn't perceived. This suggests a possible bias against newcomers and unregular contributor, with an emphasis on women [10].

2.3 Competence-Confidence Gap

The competence-confidence gap describes the phenomenon, where womens competence doesn't directly translate to confidence. As an additional variable women require the support of the community - this is dubbed as "social attraction". Even very accomplished women might suffer from it, Wang et al. name the Facebook COO Sheryl Sandberg as one example. Through this lack of confidence they lose many opportunities, as confident people are more likely to be hired or might be more reluctant in taking chances [12].

Wang et al. hypothesized women will increase in contributions once their follower count increases, insinuating that event will give them the necessary push in confidence. For that they sampled the top female contributor on GitHub and the same amount of randomly selected male contributor. They were able to prove that a correlation between surge in follower and opening pull-requests exist for women [12].

2.4 Descriptive and Prescriptive Stereotypes as Barriers

Descriptive stereotypes define how women and men are perceived to be, while prescriptive stereotypes describe how they are expected to be. Such descriptive stereotypes associate communal domains with women and agentic domains with men. Prescriptive stereotypes act in a similar way as women are expected to be caring and men are expected to be assertive. What's even more troubling is that the prescriptive stereotypes also define how women and men are not expected to be [3].

According to an experiment conducted by Rudman et al. showing assertiveness during the hiring process is disadvantageous for women. Furthermore in their study, the only applicant who didn't experience bias was one exhibiting both traits: kindness and expertise. Rudman et al. have used the analogy of this being like walking on a tightrope, as women have to compensate their agency with communal traits [8].

Imtiaz et al. go a step further and hypothesize that this tightrope of socially-accepted ways of conducting oneself influence how women interact on GitHub. They were able to prove their hypothesis where women, amongst other things, avoided overly emotional reactions and swear words [4].

2.5 Social and Cultural Barriers for Newcomers

Steinmacher et al. defined six categories of barriers newcomers face when they desire to contribute to a OSS project: they deal with the responses of other contributors, newcomers' characteristics, onboarding newcomers, issues with the documentation, cultural differences

and technical problems. Four categories out of them are social/cultural issues and thus shall be explained further [9]:

- The issues found with the responses of the other contributors dealt with the way collaborator replied: the messages were either lacking, delayed or rude. However, some of them were also too complex and could only be understood by veterans of the field.

- The newcomers themselves can be a barrier, as they often start contributing before completely knowing the process or having enough knowledge about the technology. However, they can also behave impatient or lack engagement.

- For newcomers it might at first be overwhelming. Due to missing documentation or lack of a mentoring system they often don't know where or how to start.

- With the collaboration on OSS projects newcomers can meet people from all over the world. With this, however, different cultures can clash. Messages, that are meant to be honest and direct, can in a different culture be considered rude [9].

2.6 Tools and Infrastructure Barriers for Female Newcomers

As tools and infrastructure barriers Mendez et al. called technological barriers newcomers face. Inadequate documentation and technical problems, as observed in the study conducted by Steinmacher et al., are part of tools and infrastructure barriers [5][9]. Mendez et al. hypothesize these issues to be a reason why newcomers are quick to drop-out of OSS projects [5].

The process they have used in analysing their collected data is called GenderMap - the Gender Inclusiveness Magnifier. The theory behind this is that women and men have a tendency on what kind of approach they use to solve a problem. Consequently, when a tool isn't supporting approaches associated with both genders it can be a disadvantage for either women or men. The process builds on four key-differences [5]:

- Women are more likely to use technology in achieving a goal, while men simply use it because they like it

- Women tend to have lower self-esteem in their computer-skills

- Women are more likely to avoid risks

- Women and men process information differently: women primarily gather all available information, while men take the first possible path. Furthermore, women are less likely to learn new features through trying them out; they learn these features through a process [5]

In their study they use a diverse setup of team consisting of mixed teams and teams consisting only of one gender. These teams had to work through certain use-cases, some examples being „Use GitHub issue tracker to find an issue" or „Find help with pull requests on GitHub". The problems the different teams ran into varied from unclear terminology in the documentation to missing information or lacking guides on how to contribute. In total, 73% of the found barrier-types don't support methods women are likely to use and thus are biased against them [5].

2.7 Conclusion

The previous section focused on issues and barriers women commonly face on GitHub. It's notable that women are more likely to work on the documentation of a project, than implementing new features. In addition to that, women have little confidence in their skills. These are some of the issues found, however, a common factor is how newcomers face many struggles in their first experiences with GitHub and women could be easier discouraged by them.

3 Possible Solutions for GitHub

3.1 Supporting Female Newcomer through an Onboarding-Process

One crucial issue mentioned in the previous section is how newcomers feel overwhelmed with, amongst other things, the processes, technology and terminology. These are likely reasons why they stop contributing at a very early stage. Most of these barriers can be overcome with a proper onboarding process. One approach could be assigning mentors to them. These mentors should be veterans of the field or at least regular contributor, who are familiar enough with the workflow. Assigning mentors should be the project owners responsibility. They might have a good grip on possible candidates. It is crucial that this onboarding should tackle the social and technical barriers. Making a guide, which should be short and precise, is one way of solving some technical barriers. It would at least make them familiar with the terminology. However, there's only so much that could be explained and some issues just come up on first-hand experience. That's why the newcomer should work together on their first task with the mentor. Through this the mentor could support them in crucial, however necessary, issues, like setup of the local environment, or where guides and guidelines can be found.

The previously mentioned method could be used for women and men. However, the low participation of women need to be considered separately as well. The owner of the repository need to be aware of this problem and recruit able women as mentors. This woman could be regarded as a role model by the newcomer. The lack of female role models is hypothesized to be a reason why women avoid the computer science field [2]. Working on an issue collaboratively has been shown to be beneficial for women [13]. While in that study the focus was on pair-programming, what the proposed onboarding-strategy and pair-programming have in common is the collaborative work on an issue. However, that problem could be resolved with the introduction of videoconferences.

3.2 Documentation and Guidelines

As inadequate documentation is a reason why newcomers feel overwhelmed, once they join a new project, there should be more effort put into keeping it up-to-date. This part, as well, could be part of the onboarding. In addition, to the mentor and newcomer working on their first programming task together, they could also focus on providing a proper documentation for it. This documentation can be in the form of well-formulated pull-requests and inline comments in the code, where necessary. Furthermore, a well-structured code with well-defined names for the files, classes, methods and fields could also work as documentation.

On another note, the mentor and newcomer should also work on a documentation-task together. This method would have similar benefits to the collaborative work on a coding-task, where the newcomer gets familiar with the structure for the documentation.

Furthermore, through the questions and input of the newcomer the mentor quickly becomes aware of the places where the documentation is lacking.

When resources for possible mentors are scarce, the owner and collaborator could write guidelines for newcomers. In these guidelines they should write the most important things that should be considered when one wants to contribute to their project. This set of guidelines should appear to the newcomer the moment they first want to open a pull-request in a project. Furthermore, this message should contain some kind of contact person, that the newcomer could write in case of open questions. An additional bonus would be if the contact person would additionally message the newcomer, as shy people would be more likely to talk about their issues once asked.

4   Conclusion

In this paper, the current state-of-the-art regarding gender gap on GitHub was reviewed. On that platform women are sparsely represented. The amount of women owning repositories or being very active members is low. One major problem is the confidence-competence gap of women.

It is suspected the reason for the low participation is how women still suffer under prescriptive and descriptive stereotypes. In one study it was found women avoid overreaction or swear words in their comments. Furthermore, the conversation culture on GitHub leaves room for improvement. Newcomers are faced with impolite, delayed or no responses to their queries. Aside from that, newcomers themselves often start collaborating before getting familiar with the procedures or technology. Using GenderMap, Mendez et al. have concluded that GitHub lacks support in the preferred problem-solving approach of women.

Finally, as a possible solution an onboarding-process, which would ideally assign mentors to newcomers is proposed. For female newcomers using women as mentors might lower the dropout rate. Furthermore, the same could be achieved through keeping the documentation up-to-date and clean and provide guides for newcomers with a contact person.

References

[1] Education in austria. https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/soziales/gender-statistik/bildung/index.html. Accessed: 2020-02-07.

[2] Janet Carter and Tony Jenkins. Where have all the girls gone? what entices female students to apply for computer science degrees. 2001.

[3] Madeline Heilman. Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32:113–135, 12 2012.

[4] Nasif Imtiaz, Justin Middleton, Joymallya Chakraborty, Neill Robson, Gina Bai, and Emerson Murphy-Hill. Investigating the effects of gender bias on github. In *Proceedings of the 41st International Conference on Software Engineering*, ICSE '19, page 700–711. IEEE Press, 2019.

[5] Christopher Mendez, Margaret Burnett, Hema Padala, Zoe Steine-Hanson, Claudia Hilderbrand, Amber Horvath, Charles Hill, Logan Simpson, Nupoor Patil, and Anita

Sarma. Open source barriers to entry, revisited: a sociotechnical perspective. pages 1004–1015, 05 2018.

[6] Marco Ortu, Giuseppe Destefanis, Steve Counsell, Stephen Swift, Roberto Tonelli, and Michele Marchesi. How diverse is your team? investigating gender and nationality diversity in github teams. *PeerJ*, 07 2016.

[7] Gregorio Robles, Laura Reina, Jesus Gonzalez-Barahona, and Santiago Domínguez. Women in free/libre/open source software: The situation in the 2010s. pages 163–173, 05 2016.

[8] Laurie Rudman and Peter Glick. Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, 57:743 – 762, 12 2002.

[9] Igor Steinmacher, Tayana Conte, David Redmiles, and Marco Aurelio Gerosa. Social barriers faced by newcomers placing their first contribution in open source software projects. 03 2015.

[10] Josh Terrell, Andrew Kofink, Justin Middleton, Clarissa Rainear, Emerson Murphy-Hill, Chris Parnin, and Jon Stallings. Gender differences and bias in open source: Pull request acceptance of women versus men. *PeerJ Computer Science*, 3:e111, 05 2017.

[11] Bogdan Vasilescu, Daryl Posnett, Baishakhi Ray, Mark G.J. van den Brand, Alexander Serebrenik, Premkumar Devanbu, and Vladimir Filkov. Gender and tenure diversity in github teams. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, page 3789–3798, New York, NY, USA, 2015. Association for Computing Machinery.

[12] Zhendong Wang, Yi Wang, and David Redmiles. Competence-confidence gap: A threat to female developers' contribution on github. In *Proceedings of the 40th International Conference on Software Engineering: Software Engineering in Society*, ICSE-SEIS '18, page 81–90, New York, NY, USA, 2018. Association for Computing Machinery.

[13] Linda Werner, Brian Hanks, and Charles Mcdowell. Pair-programming helps female computer science students. *ACM Journal of Educational Resources in Computing*, 4, 03 2004.