

Article

Barriers Faced by Women in Software Development Projects

Edna Dias Canedo ^{1,*}, Heloise Acco Tives ², Madianita Bogo Marioti ², Fabiano Fagundes ²
and José Antonio Siqueira de Cerqueira ¹

¹ Department of Computer Science, University of Brasília (UnB), P.O. Box 4466, Brasília, Brazil; jose.siqueira.cerqueira@gmail.com

² Department of Computer Science, Lutheran University Center of Palmas (CEULP), P.O. Box 85, Palmas, Brazil; heloise.acco@gmail.com (H.A.T.); madianitab@gmail.com (M.B.M.); thilfa@gmail.com (F.F.)

* Correspondence: ednacanedo@unb.br; Tel.: +55-061-98114-0478

Received: 2 September 2019; Accepted: 30 September 2019; Published: 9 October 2019



Abstract: Computer science is a predominantly male field of study. Women face barriers while trying to insert themselves in the study of computer science. Those barriers extend to when women are exposed to the professional area of computer science. Despite decades of social fights for gender equity in Science, Technology, Engineering, and Mathematics (STEM) education and in computer science in general, few women participate in computer science, and some of the reasons include gender bias and lack of support for women when choosing a computer science career. Open source software development has been increasingly used by companies seeking the competitive advantages gained by team diversity. This diversification of the characteristics of team members includes, for example, the age of the participants, the level of experience, education and knowledge in the area, and their gender. In open source software projects women are underrepresented and a series of biases are involved in their participation. This paper conducts a systematic literature review with the objective of finding factors that could assist in increasing women's interest in contributing to open source communities and software development projects. The main contributions of this paper are: (i) identification of factors that cause women's lack of interest (engagement), (ii) possible solutions to increase the engagement of this public, (iii) to outline the profile of professional women who are participating in open source software projects and software development projects. The main findings of this research reveal that women are underrepresented in software development projects and in open source software projects. They represent less than 10% of the total developers and the main causes of this underrepresentation may be associated with their workplace conditions, which reflect male gender bias.

Keywords: gender diversity; open-source projects; software development projects; women in STEM; systematic literature review

1. Introduction

Currently, the global economy is highly dependent on software, and it is becoming more common for companies that do not have software development as their end product to carrying out this activity to support their business [1]. Organizations, in general, are structured to meet the needs of the market and to adapt to the innovations that occur in the area of technology.

The closed development model, in which a business owns the intellectual property of a software and is fully responsible for its development, was the predominant model for the software development

industry in the early 2000s [1]. However, this predominance did not last long, since in a short period of time the open source software development model (OSS), “where the company shares the software development code” [2], became part of the routine of software development companies.

Major companies in the information technology industry are opting to open source their strategic products in order to gain a competitive advantage. Among them, Amazon, Google, Facebook and Microsoft have embraced open source for their artificial intelligence solutions. The decision to open source for this kind of software—which can give a considerable competitive advantage to a company—may not seem to promise a financial return, however, what these companies are likely to expect as a return is the opportunity to be the basis for future innovation [3].

Multiple studies show that women are underrepresented in almost all fields of Science, Technology, Engineering and Maths (STEM) [4]. This gender gap is also present at higher education institutions in both student numbers and academic staff [5]. There has been an increasing number in studies approaching women in a male-dominant field. Women have entered many other previously male-dominated fields, including other STEM fields, but not computer science and engineering [6]. While social capital is beneficial to the long-term engagement of both sexes, women appear in small numbers on most software development teams [7]. Therefore, the understanding of the profile of women working in open source projects and identification of factors that influence the permanence and engagement of these women in teams can assist in pinpointing problems or bottlenecks that may be hindering women’s participation.

LinkedIn’s 2018 diversity report shows that women represent 42.9% of their workforce, with a 39.1% representation in leadership positions, representing a 12% increase in the last two years [8]. Despite this increase, the representation of women in the technology area is 21.8%, while men represent 78.2%. Some organizations aim to change this underrepresentation, by adopting programs committed to a set of goals to increase women’s workforce participation and create a more inclusive culture. For an example, a recent report about diversity in Google’s company shows that hiring women has grown to 33.2% [9], an increase of more than 1.9% regarding the previous report. In particular, the concentration of women in non-tech areas increased to 47.2%, representing an increase of more than 3.3 points. Although the hiring of women has increased, hiring for companies’ leadership positions has decreased by 25.9% (−3.5 points) from the previous report [9].

In this context, this paper proposes a systematic literature review (SLR) in the main databases of scientific works in the computation field with the aim of discovering factors that can help increase the rate of women’s interest in OSS projects and software development projects.

The main findings of this research were that women are underrepresented in software development projects and open source software projects. The main causes of underrepresentation of women may be related to their workplace, as there is—implicitly—a gender bias from the part of men, towards women. In addition, women are more likely to abandon their careers in technology. We have not found in the literature a model for increasing women’s engagement in software development projects. An analysis of existing initiatives to increase women’s participation in STEM needs to be undertaken, as studies show that current solutions are not having a positive effect. Increasing the number of women in STEM may lead to an increase in the number of women in open source software projects and software development projects.

We see varying representation from men and women in different developer roles on our survey. All categories have dramatically more developers who identify as men than women but the ratio of men to women varies. Developer types above the line have respondents that are more likely than average to be men, and those below the line have respondents who are more likely than average to be women. Developers who are data scientists or academic researchers are about 10 times more likely to be men than women, while developers who are system admins or software development and information technology operations (DevOps) specialists are 25–30 times more likely to be men than women. Women have the

highest representation as front-end developers, designers, data scientists, data analysts, quality assurance (QA) or test developers, scientists, and educators.

The remainder of this paper is organized as follows: Section 2 presents the relevant concepts to support the research and related works. Section 3 presents the conduction of the systematic literature review. In Section 4, we present the results obtained from the systematic literature review. Section 5 presents an analysis and discussion of the results, as well as limitations and threats to the validity, and lastly, in Section 6 we present our conclusion and future works.

2. Background

According to UNESCO [10] gender equality exists when women and men enjoy the same status and have equal conditions, treatment, and opportunities to achieve their full potential, human rights, and when they contribute to and benefit from economic, social, cultural and political development [10]. Gender diversity deals with equal representation of men and women in the workplace.

Research has shown that diversified software development teams are a factor that is directly associated with increased productivity in projects [11,12]. A diverse team can be described as a composition of members of different ages, with different levels of experience and knowledge in the area, as well as the participants' gender.

A more diverse software development team is closer to the end users of a software, as these users are typically considerably diversified. Therefore, a more diversified development team has a greater possibility to understand and represent the end user's needs, contributing to an appropriate alignment between the software produced and its possible users [13].

Team diversity is a factor that can be explored to a great degree in open source software development projects, given that a decentralized team working with the common goal of creating open source software is a characteristic of this kind of project. Generally in this type of environment, a considerable variety of culture, language and habits is observed from the participants, who often are from geographically separated localities [11].

The scope of the open source definition exceeds the plain access to the source code of the product/system to be developed. According to the Open Source Initiative (OSI), it is a set of criteria to be followed for software distribution, such as [2]:

1. Free redistribution: The license must not have any restricted part to be sold or distributed separately, such as a software distribution component composed of several different codes. The license will not require royalties or other fees for the sale;
2. Source code: The compiled form and the source code must be included in the program that will be distributed. If the product is not distributed with the source code, there should be a well-known manner to obtain the source code at a reasonable cost to be downloaded over the internet or reproduced. The code must be readable and intelligible so that any programmer can modify the program;
3. Derived works: The license must allow modifications and derived works. The products originated from modifications must be distributed under the same license terms from the original software;
4. Integrity of the source code of the author: The license must explicitly allow the distribution of the program built from the modified source code or require that derived programs have a different name or version number of the original software;
5. Non-discrimination against persons or groups: The license can not be discriminatory against any person or group of persons;
6. Non-discrimination against areas of practice: The license should not restrict anyone from using the program in a specific field of activity;

7. Distribution of the license: The rights associated with the program must apply to all those whose program is redistributed, without the need to execute an additional license for these parties;
8. Non-specific license for a product: The rights associated with the program must not depend on the program being part of a specific software distribution. If the program is extracted from this distribution and used or distributed within the terms of the program license, all parties for whom the program is redistributed must have the same rights as those granted in connection with the distribution of the original software;
9. License not restricted to other software: The license must not contain restrictions to other software that are distributed with the licensed software;
10. Technology-neutral license: No license clause can establish an individual technology or interface style to be applied in the program.

When software products meet the distribution criteria and are classified as open source software, a number of advantages can be obtained, such as: An increase of creativity in development, improvements in product quality, a decrease in production cost, and faster identification of failures [14].

Related Works

The latest studies show that the number of female students enrolled in science, technology, engineering and mathematics (STEM) related disciplines have been decreasing over the last twenty years. Unfortunately, the number of women resigning from technological job positions also remains unacceptably high [15–17]. Furthermore, recent diversity reports demonstrate a large gap between the percentage of women holding computing jobs compared to the percentage of men [18].

In a survey conducted by Olmedo-Torre et al. [17] with over a thousand STEM female graduate students, social stereotypes and immediate environment were the main reasons pointed out for low enrollment of women in STEM.

Conversely, Reinghans et al. [15] developed a model for retention of girls in computing and engineering fields, through the creation of a community that provides academic and professional resources to them, overcoming the stereotype threat.

Following the sense of community, Vaschovsky et al. [18] evaluated the Stanford Artificial Intelligence Laboratory's Outreach Summer (SAILORS) as a way of addressing the lack of diversity in artificial intelligence (AI). This camp is about female high school students who are willing to learn artificial intelligence and attract more women to the field of computer science. They analyzed the program and in conclusion, the results showed a positive impact from SAILORS by achieving the goals of contextualizing technical AI concepts through social impact and addressing barriers for girls in computer science.

Another study that makes use of innovation to attract women in STEM is presented by Botella et al. [16]. They argue that one action to reduce the gender gap is the gendered innovation initiative, and in particular, Machine Learning and Data Science areas provide new opportunities to include gendered innovation in Information Theory, which could be achieved by the fact that these areas are very diverse.

GitHub conducted a survey to collect data on open-source software development practices. The survey collected responses from 5500 contributors from more than 3800 open source repositories in GitHub and more than 500 responses from community contributors working on other platforms. The results showed that only 3.36% of survey contributors were women, while 90.95% of contributors were men [19]. In addition to the problem of under-representation in the open-source community, when women identify their gender in their profiles, their acceptance rates for pull requests are lower than when their gender is not identifiable through their profile [20].

Ford et al. [21] interviewed 22 women (students and practitioners) to understand how they used Stack Overflow. The study identified 14 barriers that affected women and led them not to contribute to the Stack Overflow community. Furthermore, the authors concluded that women are more likely to contribute when they find other women contributing in the same community [22].

In the work developed by May et al. [23] the gender differences in activities and user success in Stack Overflow were investigated. Women are highly underrepresented, not only in comparison to male users, but also in relation to their presence in the Information Technology job market. Furthermore, according to the success measures used by Stack Overflow (the reputation of the site is given from the activities of users) women are less successful than men, since the latter is more active in their activities (men answer questions more than women, and therefore have a higher reputation), although women ask many questions. The authors propose a new reward mechanism to calculate the reputation of the contributor that does not favor only male behavior.

3. Systematic Literature Review

This work presents the conduction of a systematic literature review (SLR). An SLR is a way of identifying, analyzing and interpreting available evidence related to a particular research question, area or phenomenon of interest [24,25]. Studies that contribute to a SLR are called primary studies. A systematic review is considered a form of secondary study. During the execution of this work we followed the phases of Planning, Conduction, and Results [24,26,27].

- Planning: Aims to identify the real need of the SLR, in other words, the motivation for the execution of research [25]. This phase is composed of the main activities: to define the objective, to prepare the protocol that will guide the SLR in order to minimize biases that can be committed by the researchers, and to evaluate this protocol, with the execution of the test of the research protocol in the data bases;
- Conduction: Performs the application of the search strategy in order to identify and select studies according to the protocol defined in the planning phase. From the set of selected studies the data necessary to compose the results of the work must be extracted and synthesized [25];
- Publication of Results: Prepares the final SLR documentation, containing the description of the results and the answers to the research questions defined in the work protocol. The results, where possible, should be disclosed to potential participants [26].

The tool (State of the Art through Systematic Review—StArt) [28] auxiliated in the planning and conduction phases of the systematic literature review. The SLR was carried out with the objective of mapping the problems causing women's lack of interest and possible solutions to these problems. Also, to identify the profile of women who work on open source software development projects and not open source projects.

3.1. Planning of the SLR

Table 1 presents the research questions (RQ) that will comprise this systematic literature review.

Table 1. Research questions (RQ).

RQ	Description
RQ.1	What are the possible problems causing women's lack of interest (engagement) in the area of software development?
RQ.2	What are the possible solutions to increase the engagement of women in open source software development projects?
RQ.3	What is the profile of the women professionals involved in open software development projects?

3.2. Research Strategy

The search strategy involved the use of Automatic Search [29], which consists of searching through data bases with a search string, followed by Manual Search [29], in which papers in annals of conferences, congress or specific journals were searched. Automatic Search was performed in three data bases selected for having a considerable high volume of papers published in journals and conferences in the area of Information Technology and Communication, those are:

- ACM Digital Library;
- IEEE Xplore Digital Library;
- DBLP-Computer Science Bibliography.

3.3. Selection Criteria (Inclusion and Exclusion)

The following selection criteria were defined for the selection of primary studies:

- The work must be available in the previously defined digital library;
- The study must have been written in English or Portuguese and published between 2007 and 2019;
- The work must be related to the area of Information Technology or Computing;
- The study must be related to open source software projects and gender diversity or software development project and gender diversity;
- The work can be classified as gray literature, namely, technical reports, preliminary studies, technical specifications, official documents of specific organizations [25].

As the criteria of exclusion for the studies, the non-fulfillment of some of the inclusion criteria was considered, as well as:

- Incomplete works (published as a short paper);
- Works that do not present enough information to extract the expected data, thus impairing the quality or relevance of the work [30].

Regarding the evaluation of the identified studies quality—after the execution of the search strategy—the selection of the most relevant papers that compose the SLR must be conducted through the four steps of study selection, defined by [29], which are shown in Figure 1.

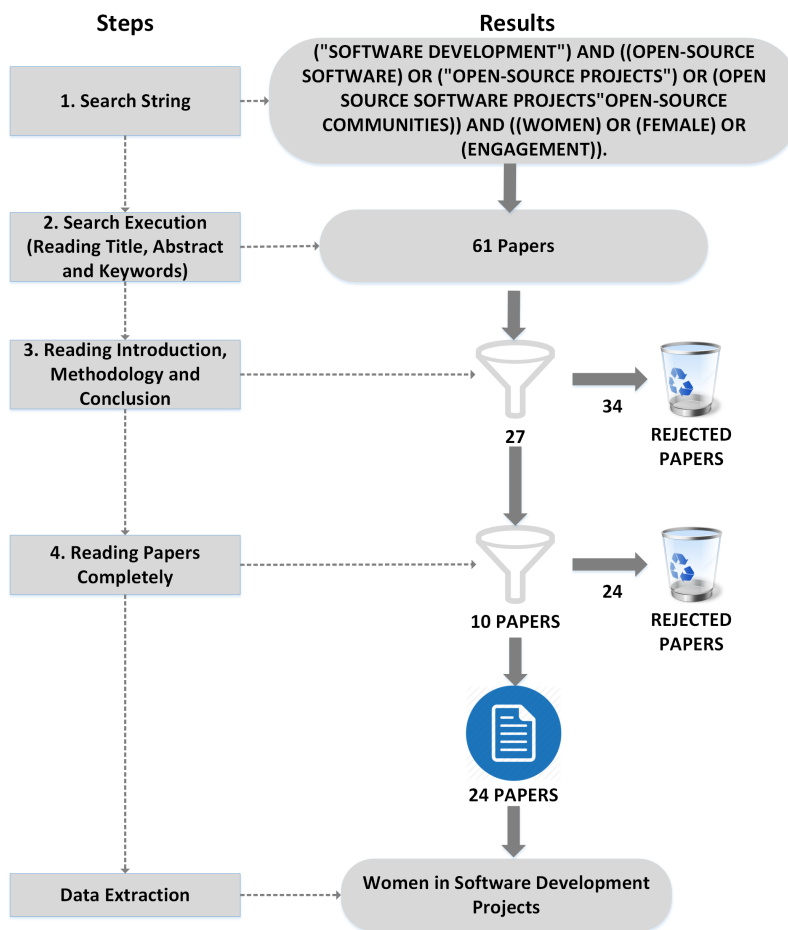


Figure 1. Steps of the selection of works strategy (Adapted from [29]).

- Step 1: Execution of the search strategy involving the automatic search. A preliminary list of studies should be generated for each data base used, and the reference file (in the .bib extension) should be downloaded to enable SLR management with the aid of the StArt tool. Duplicate entries should be deleted by the tool, while importing the file;
- Step 2: Identification of potentially relevant studies, based on reading the title and abstract. In this step, it should be possible to rule out studies that are clearly irrelevant to the research. In case of doubt about the permanence of some study in the SLR, the next step may assist in this definition;
- Step 3: Reading of the introduction, methodology and conclusion of the pre-selected works, again applying the inclusion and exclusion criteria;
- Step 4: The papers selected in Step 3 should be read in full and the volume of papers resulting from this step should be used to compose the SLR and support the answers to the research questions.

3.4. Conduction of the SLR

The process of the automatic search of the primary studies, consisting of Step 1 of the selection strategy, was executed with the string defined in the protocol in the data bases previously defined. The automatic search on the three data bases resulted in a total of 61 papers, of which 29 or 47.5% of the articles came from the <http://ieeexplore.ieee.org> (IEEE Digital Library), 19 or 31.2% of the articles from the <http://dl.acm.org/> (ACM Digital Library) and 13 or 21.3% of the works were located at <http://dblp.uni-trier.de/> (DBLP-Computer Science Bibliography). It is important to note that five studies were identified as duplicates and have already been eliminated during the initial load of the StArt tool.

The execution of Step 2 of the selection strategy, which consisted of reading the title and abstract of the papers, reduced the volume of pre-selected papers to 34 articles. During this step, the information required in the StArt tool was totally filled, since IEEE Digital Library alone, in its bibtex file, brought in all of its required data (Author, Title, Keywords, Journal, Abstract and Year).

After the execution of the fourth and last step of the selection strategy, which involved the complete reading of the articles, 24 primary studies were selected to be used in the data extraction.

The progress of resulting works at each step of the article selection strategy accumulated per year of publication is presented in Figure 2.

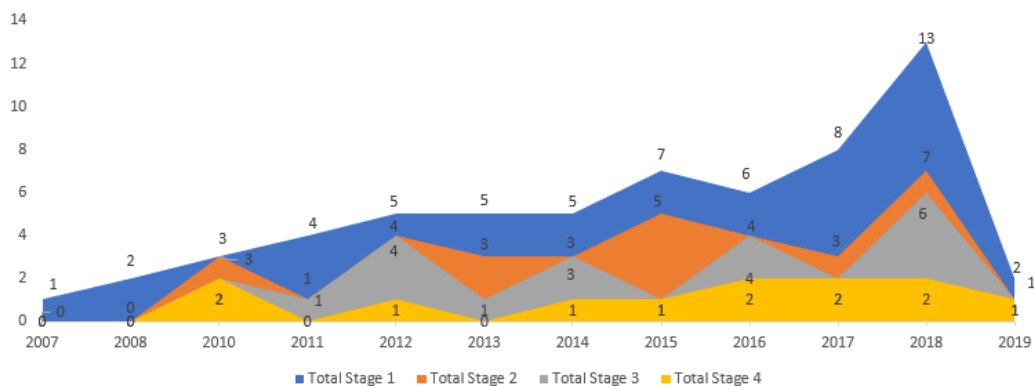


Figure 2. Selection progress of works per year.

In Figure 2, it is possible to identify that the number of publications involving the participation of women in open source projects is in a phase of growing interest by researchers in the last 10 years, with an increase in the number of publications in the last 2 years.

3.5. Data Extraction

The primary studies selected in the SLR are presented in Table 2.

Table 2 is composed of the following fields: Primary Study, column identifying the selected studies; Reference, to present identification of the author of each work; and RQ, information on which research question was answered with the help of the identified work.

Table 2. Relation of selected articles.

Primary Study	Title	Reference	RQ
S1	A study of FOSS 2013 survey data using clustering techniques	[31]	2, 3
S2	Code of conduct in open source projects	[32]	2
S3	Gender and Tenure Diversity in GitHub Teams	[11]	3
S4	Gender differences and bias in open source: pull request acceptance of women versus men	[20]	1, 3
S5	Gender differences in early free and open source software Joining Process	[33]	
S6	Gendered behavior as a disadvantage in open source software development	[34]	1, 2
S7	Gendered patterns of politeness in free/libre open source software development	[35]	
S8	Joining and Socialization in Open Source Women’s Groups: An Exploratory Study of KDE-Women	[36]	2
S9	Making a Difference: An Overview of Humanitarian Free Open Source Systems	[37]	2
S10	On older adults in free/open source software: reflections of contributors and community leaders	[38]	1
S11	Open Source Barriers to Entry, Revisited: A Socio-technical Perspective	[39]	1, 2, 3
S12	OpenStack Gender Diversity Report	[13]	1
S13	Patches don’t have gender: What is not open in open source software	[40]	
S14	Some of All Human Knowledge: Gender and Participation in Peer Production	[41]	2
S15	Women contributions to open source software innovation: A social constructivist perspective	[42]	
S16	Gender diversity in STEM disciplines: A multiple factor problem	[16]	1, 2
S17	A model for increasing gender diversity in technology	[15]	1, 2, 3
S18	Do female motives for enrolling vary according to stem profile?	[17]	1, 2, 3
S19	Toward more gender diversity in CS through an artificial intelligence summer program for high school girls	[18]	2
S20	FLOSS Participants’ Perceptions about Gender and Inclusiveness: A Survey	[43]	1, 3
S21	Why are some STEM fields more gender balanced than others?	[44]	2, 3
S22	Diversity and Decorum in Open Source Communities	[45]	2
S23	Gender in Open Source Software: What the Tools Tell	[12]	2
S24	How Long and How Much: What to Expect from Summer of Code Participants	[46]	1, 2, 3

4. SLR Results Synthesis

When applying the automatic search strategy, we identified a total of 61 articles, which were analyzed during the steps defined in Figure 1. After the execution of all the steps of the selection strategy, a total of 24 papers were selected and already presented in Table 2. These works were used to answer the research questions, as presented below.

4.1. RQ.1. *What Are the Possible Problems Causing Women's Lack of Interest (Engagement) in the Area of Software Development?*

An already mapped fact of open software development projects is prejudice and discrimination based on the gender of the participant. To be considered a woman is not only a disadvantage for women, but also for men who do not define their gender together with members of the project [34]. This disadvantage is given in relation to the rejection considered natural with this group of participants, as well as the denial of requests for participation in the project and the contributions made.

When it comes to gender, the field of software engineering is strongly male orientated, this is evidenced through several studies that show that the gender situation in the area of open software development projects is even more unbalanced [13]. A 2013 survey of more than 2000 open source developers indicated that only 11.2% were women [47]. Despite efforts from both the academy and the industry, the situation has not shown significant changes over the past 15 years, resulting in what has been called a lost decade in the integration of women in the field of open source [13].

A study conducted by [38] indicated that 63.63% of male members of open software development projects reported having experienced or observed gender discrimination. This fact was also studied by [42], where it was identified the need to apply research and studies not only in the identification of the gender of the participants of open software development projects, but also in the social constructions among the participants.

The tools and the infrastructure of open software development projects, which are used by team members to access project resources, interact with other members, and learn about the contribution process, have been identified as factors that can disrupt the integration of women by generating either disengagement or lack of engagement in projects.

This finding was the result of the research of [39], where it was verified that the majority of women who are beginners in open software development projects, have greater difficulty of adaptation to the tools and infrastructure available than the men, and usually end up considering their performance unsatisfactory.

4.2. RQ.2. *What Are the Possible Solutions to Increase the Engagement of Women in Open Source Software Development Projects?*

It is essential that open source communities protect their members against prejudicial, homophobic, exclusionary or destructive behaviors and provide a welcoming, secure, friendly and inclusive environment in which people can effectively collaborate to present successful products [32].

Groups interested in expanding women's participation in open software development projects have been developing a series of effective strategies to improve the work environment [41]. These initiatives have been adopted both in academia and in the industry in open software development projects. Some of the good practices adopted in this context are presented in Table 3.

Recommendations in reports and studies aimed at increasing the productivity and quality of software and systems developed from open software development projects point out that team diversity is a beneficial factor in the success of projects [13]. Thus, the use of inclusive actions of women in projects through the practices identified and the greater understanding of their profile are increasingly important for the success of the projects.

Table 3. Good practices adopted.

ID	Good Practices	Reference
1	Setting quotas for women has become a usual practice. The precaution that must be taken into account regarding this type of action is the need to observe that there is no decrease of feminine behavior characteristics and stereotypes are not created on the competence of this group of professionals.	[34]
2	Initiatives to assist in the development of leadership skills and technical performance of women can be adopted.	[36]
3	Transparency regarding the organizational culture where the project is inserted as well as the adopted work policy. This type of attitude provides conditions for members to identify characteristics of the project's work environment even before deciding to participate in the team. It is expected that this will increase the degree of affinity between personal and professional values.	[36]
4	The adoption of a code of conduct in open software projects is an emerging phenomenon to deal with diversity issues and provide a more secure and inclusive community. A code of conduct simply establishes basic rules for communication between participants, outlines mechanisms and attempts to code the spirit of a community, in a way that any person may contribute comfortably, regardless, for instance, gender, ethnicity or sexual orientation. An active awareness of community members about their code of conduct is essential to the effectiveness of its application. Behavioral guidelines should be emphasized as well as approaches in case of violation, which can range from a period of time to take appropriate corrective action, or even the banishment of certain project members.	[32]
5	Anti-harassment policies and active moderators in communities assessing existing interactions can also be used.	[41]
6	Women's support groups that serve as a support network for marginalized participants can be implemented.	[41]
7	Projects that are inserted in the context of Humanitarian Free Open-Source Software (HFOSS) and which have philanthropic objectives, due to their altruistic nature, have been seen as a way, not yet proven, to engage women developers in open software development projects.	[37]
8	The need to learn new skills has been pointed out as a motivator for women to enter, engage and persist in the world of open software development projects.	[31]
9	Continuous monitoring of female participation can be used to generate metrics. Such data can be used to achieve a better understanding of the female profile, identifying ways to improve collaboration with project team leaders and other community leaders. Generated documentation can facilitate the publication and sharing of best practices for the ongoing support of women in working groups.	[13]

4.3. RQ.3. What Is the Profile of the Women Professionals Involved in Open Software Development Projects?

Research conducted by [11] with GitHub users, pointed out that team members consider that the commitment and the need to perform different activities in the project (such as: push code, submit pull requests, report issues) is feasible to every member in the project, without a clear distinction of the role to each one [11], independent of any characteristic of the team member. Moreover, the research carried out by [13] tried to verify, more specifically, the leadership roles of women in open software development projects. The results are presented in Table 4.

Table 4. Leadership roles of women in open software development projects [13].

Percentage of Women	Leadership Poles
7%	Project committee members
8%	Project team leader
9%	Project board directors
7%	Technical committee
8%	Working group leaders
23%	Project ambassadors

Another study in the field, this time executed by [31] with more than two thousand active developers in open source projects, found some characteristics regarding the profile of the women participating in this type of project, presented in Table 5:

Table 5. Profile of women participating in open source software projects.

Feature	Result
Relationship Type	35% of the female participants were single, over 11% do not live with their partners, while 3% live with their partners. 3% are married and 0.09% are separated from their partners
Existence of children	20% of the women interviewed had children, while 79% did not
Current condition of employment	Around 5% of the participants were unemployed
Country of residence	40% of female participants work or reside in the US, 6% in Germany, 6% in India, 5% in the United Kingdom and 4% in Australia
Degree	81% of women who contribute are at least undergraduates. The number of participants who have a PhD is about 10% of the total
English language knowledge	82% of the contributing women are at least proficient in the English language
Job satisfaction level	32% felt they love their current job, while 5% wanted a more interesting job than the current one
Evaluation	28% consider themselves as part of the Open Source Software community

Regarding the acceptance of pull requests made by women, it was verified in [20] that they are statistically more accepted than those made by men. The women highest overall acceptance rate occurs on the account of the female survival rate within GitHub—which consists of the rate of permanence of women in projects—being much lower than that of men. Therefore, women who remain in projects tend to be more experienced, and thus better prepared to contribute than men. Other findings of the same research are:

- Women continue to have high acceptance rates as they gain experience;
- Women’s pull requests are less probable to meet an immediate need for the project;
- Changes proposed by women are broader;
- The acceptance rates of women are higher in languages such as HTML than in languages like Java or Python;
- 76% of the participating women contribute directly to coding;
- Women have lower acceptance rates when they identify themselves as women.

In summary, based on the research analyzed and already reported, it can be considered that the role of participants in open software development projects is not always clearly defined [11], but when it is, the participation of women in leadership roles rarely exceeds 10% [13].

In relation to the main characteristics of women’s profiles that participate in this type of project, it was identified that: Most of these women do not have children, most have another source of income or employment, most have at least an undergraduate degree, and most are fluent in English [31].

5. Comparison of Results with Existing Evidence

In order to analyze women’s participation in software projects, we make use of results from the Google Summer of Code (GSoC) which is a global program focused on introducing students to open source software development. Since its inception in 2005, the program has brought together more than 14,000 student participants and 24,000 mentors from over 118 countries worldwide [48]. Figure 3 presents a digest of participants in GSoC in the last three years: 2016, 2017 and 2018, organized by gender (male or female) and type of developer (mentor or student).

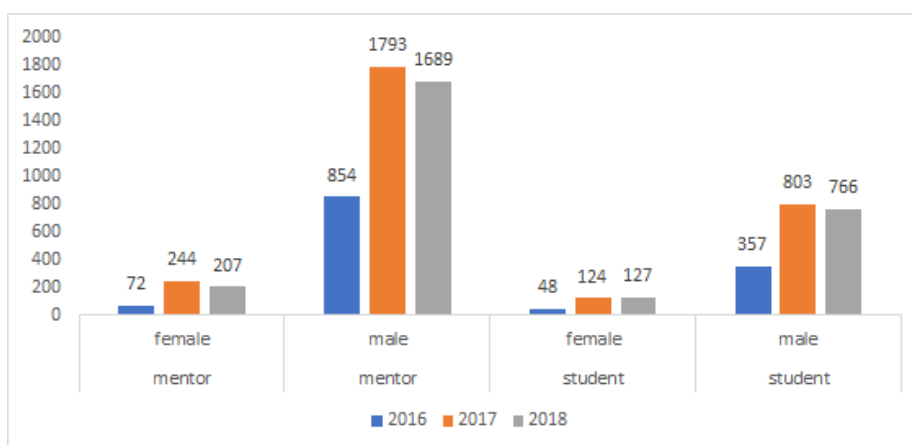


Figure 3. Google Summer of Code (GSoC) participants in the years of 2016, 2017 and 2018 [48].

Analyzing only the data regarding the gender of GSoC participants from 2016 to 2018 as shown in Figure 4, there is a minor variation between the years, but persisting the volume of women close to 11.13% of the total number of participants, which reaffirms the results of the research found in the systematic literature review articles, in which the percentage of women contributing to software projects is less than 12%.

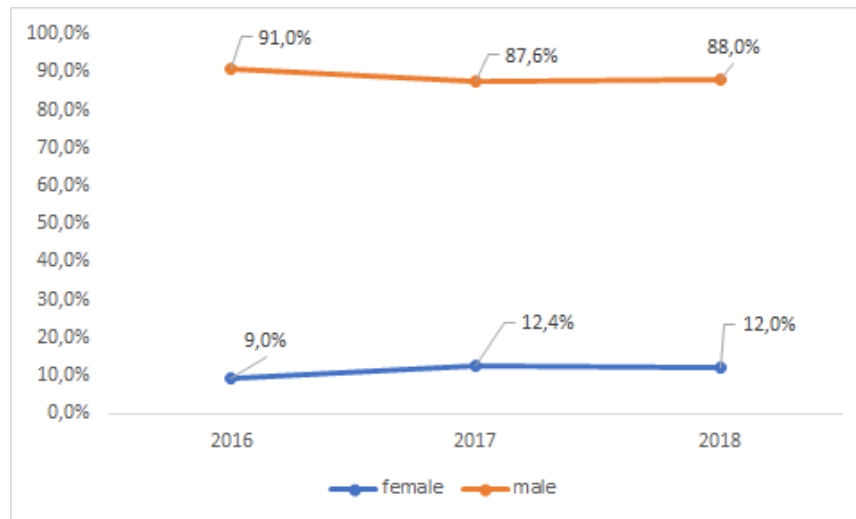


Figure 4. GSoC participants gender from 2016 to 2018 [48].

Inspecting data solely in respect to the type of developer (student or mentor) who are GSoC participants from 2016 to 2018, presented in Figure 5, it was observed that there was a minimal alteration between the years, but the volume of mentors remained close to 68% of the total number of participants in the last three editions. It is worth noting that women mentors in 2016 were 8.43% of participants, in 2017 13.60% and in 2018 12.25% of total mentors, as can be seen in Figure 3.

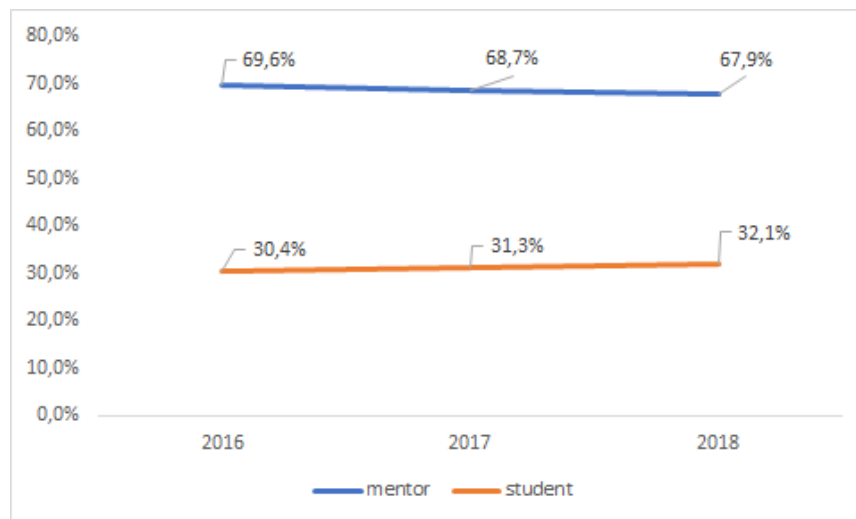


Figure 5. Developer sorted by type in GSoC from 2016 to 2018 [48].

The Stack Overflow’s annual Developer Survey is the largest and most comprehensive survey of individuals who code throughout the world. Each year, they send out a survey covering everything from developers’ preferred technologies to their job priorities [49]. From the survey results it is possible to have an understanding of the profile of women who work with software development in several countries. The percentage of women in software development is higher only when their experience is between 5 and 9 years, 34.0%. When we look at differences in years since learning to code by gender, we see evidence for the shifting demographics of coding as a profession, as well as retention problems in the tech industry for underrepresented groups. Research shows, for example, that women leave jobs in tech at higher rates than

men, as shown in Figure 6. Among our respondents, both in the United States and internationally, women are about twice as likely as men to have three years of coding experience or less. Companies interested in building a diverse developer workforce that is more reflective of society should focus on retention of their senior developers from underrepresented groups, along with thoughtful hiring from the population of more junior developers [49].

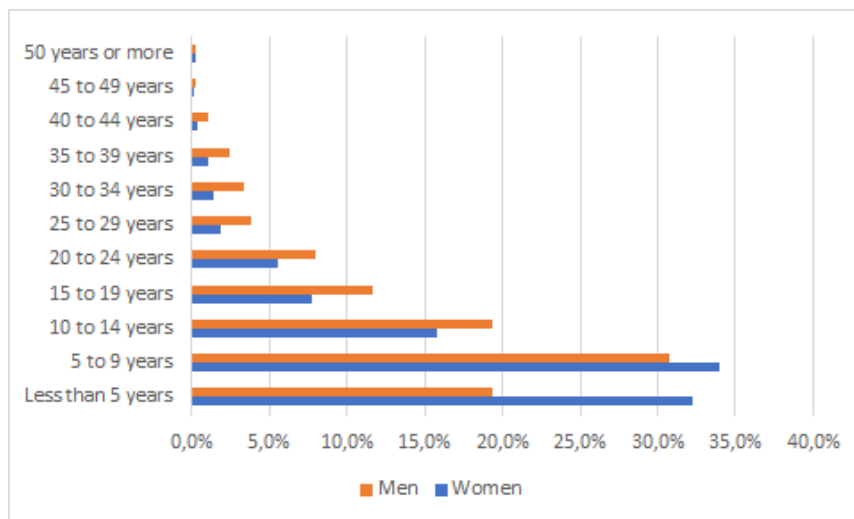


Figure 6. Experience and gender (adapted from [49]).

Figure 7 presents the results obtained from the survey regarding the gender of the developers per country. The United States ranks first with 11.7% of women acting as software developers, followed by Canada with 10.7% and the United Kingdom 8.6%. Brazil is second to last with 5.2% and lastly Italy with only 3.7% of women [49]. Overall, 91.7% of respondents are men, 7.9% are women, and 1.2% are non-binary or non-conforming gender. In addition, developers are usually young, 27.6% of participants are between 25 to 29 years old and 21.1% are between 20 to 24 years old. Regarding contributing to open source, 36.3% of survey participants never contributed, 28.1% contribute less than once a year, 23.1% less than once a month but more than once a year or 12.4% once a month or more often [49].

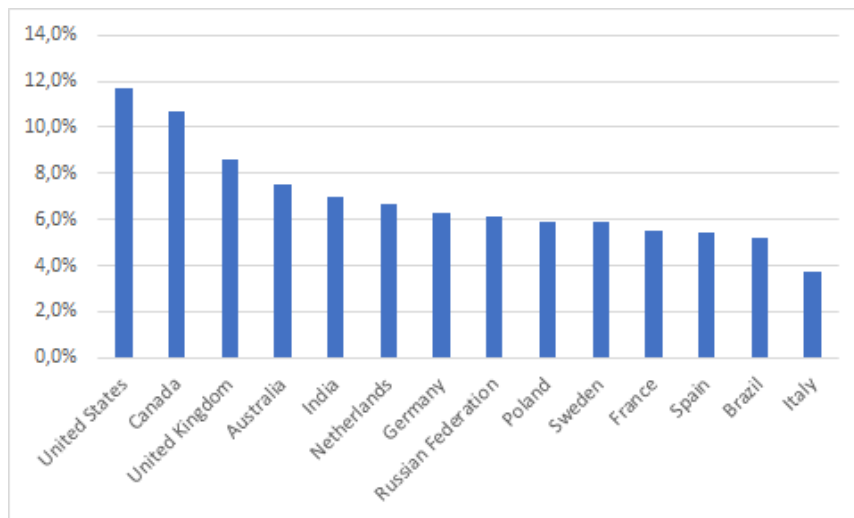


Figure 7. Gender minorities by country (adapted from [49]).

Regarding the greatest challenges to productivity, the developers responded that they are: Distracting work environment and Meetings, as shown in Figure 8. The largest difference in the relative frequency of female and male responses was in the item Being tasked with non-development work (36.9% for men and 32.8% for women). Furthermore, women consider Toxic work environment to be a more pronounced issue, 23.2%, while for men it is only 20.5%.

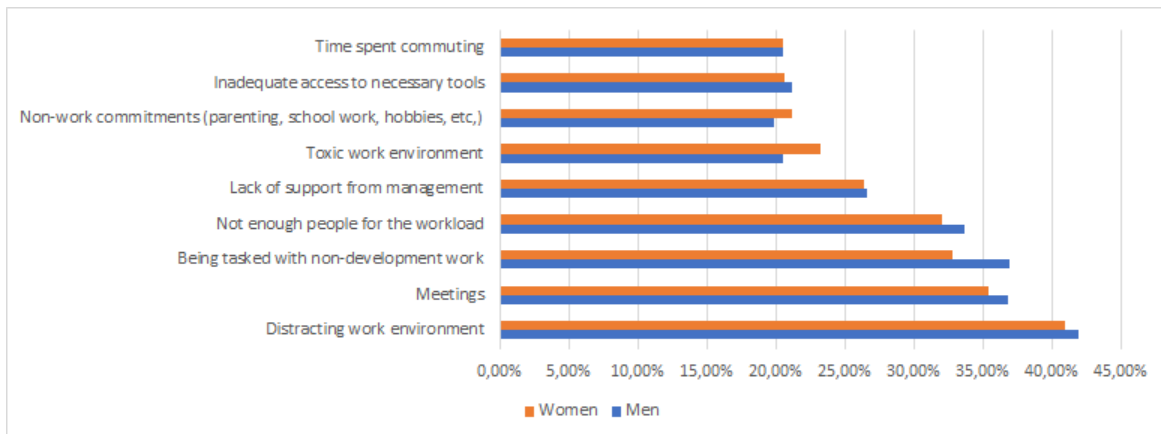


Figure 8. Greatest challenges to productivity (adapted from [49]).

The results presented from GSoC dataset extraction [48] and from Stack Overflow’s annual Developer Survey [49], are summarized in Figure 9. By means of the synthesis analysis of the data presented in Figure 9, it is possible to confirm the findings of the systematic literature review approached in Section 3, enabling a more accurate description of the scenario of women’s participation in software development projects and in open-source communities: Though increasing rates have been pointed out, women are still a target of underrepresentation in the area, reaching rates of less than 11% of the total amount of developers.

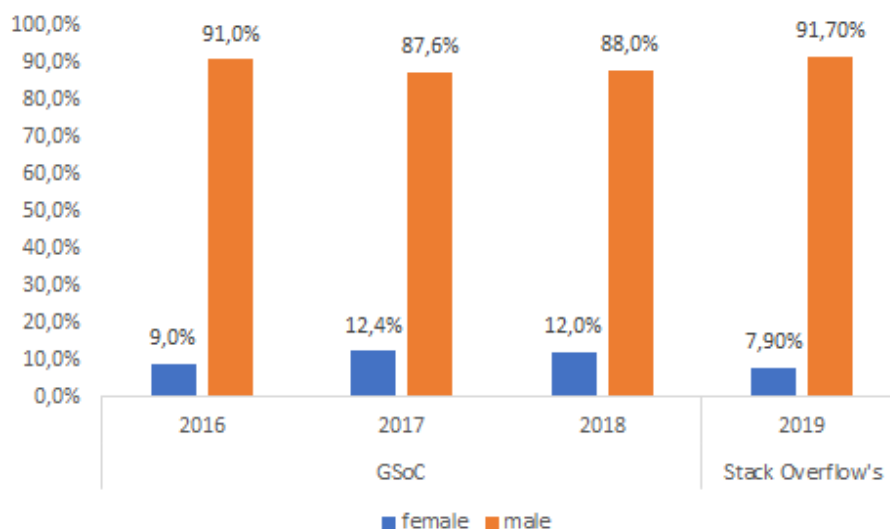


Figure 9. Number of GSoC and Stack Overflow’s participants sorted by gender (adapted from [48,49]).

Threat to Validity

We cannot guarantee that all relevant primary studies were selected. It is possible that relevant papers were not chosen. In order to mitigate this, we performed the automatic search, and complemented it by performing a manual search to try to collect all primary studies in this field. During the data extraction process, the primary studies were classified based on our judgment. In order to mitigate this threat, the classification process was performed using peer review.

Another threat to the validity of our results is the possibility that the first author might have introduced his bias in the data collection process. In this respect, the analysis process of collected data was performed along with another researcher. This researcher reviewed and analyzed all the intermediate results (primary studies and survey results). This iterative process was repeated until the end of data collection and data analysis processes. We also held meetings to validate the obtained results.

A third threat is that a literature review method based on scientific papers, in principle, may provide partial answers. To mitigate this threat we also reviewed the Stack Overflow's annual Developer Survey [49]. This survey annually investigates the preferences of code developers around the world and covers a variety of issues related to gender, personal characteristics, and techniques. The survey was answered by 90,000 developers and only 7.9% of the participants are women.

6. Conclusions

Recommendations made by research in the area of open software development projects strongly highlight the urge to increase the diversity of the participants' personal characteristics, due to the innumerable benefits that this miscellany can offer [13]. Nevertheless, engaging participants with different profiles and backgrounds in software development projects is not a simple task.

Academia and industry efforts are crucial to promote a change in the volume of female participation in this type of project, and this work, through a systematic review of literature in the main libraries of scientific works in the area of computation, sought to assist in the finding of factors that may aid in a higher rate of women's interest in open software development projects.

The main contributions of this work are in the answers of the research questions of the SLR, where it is possible to identify factors that cause women's lack of interest in open source software projects and software development participation; possible solutions were suggested to increase the engagement of this public, for instance, the effort of keeping senior developers from underrepresented groups active, along with the thoughtful hiring of more junior developers [49]. As a result of this work, it was also possible to identify the profile of women that participate in Google Summer of Code in relation to the number of students and the quantity of mentors, which corroborates with the results found in the systematic literature review where there is an underrepresentation of women, as well as Stack Overflow's annual Developer Survey results [49] where only 7.90% of software developers in the world are women.

As a follow-up to this research, we expect to deepen the study on the aspects that may influence the participation of women in open source software projects and software development projects, as well as to propose ways of addressing the identified problems regarding issues of gender inequality in open source communities and software factories.

Author Contributions: Barriers Faced by Women in Software Development Projects was made by E.D.C., H.A.T., M.B.M., F.F. and J.A.S.d.C. All authors contributed to Writing Original Draft Preparation and Writing Review and Editing.

Funding: This research work has the support of the <http://www.fap.df.gov.br/> (Research Support Foundation of the Federal District (FAPDF)) research grant 05/2018.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Milan, L.F.A.B. Adoção do Modelo Aberto de Desenvolvimento de Software Pelas Empresas. Ph.D. Thesis, Escola de Administração de Empresas de São Paulo da Fundação Getulio Vargas, São Paulo, Brazil, 2018.
2. Fitzgerald, B. The transformation of open source software. *MIS Q.* **2006**, *30*, 587–598. [[CrossRef](#)]
3. Shafto, P. Why Big Tech Companies Are Open-Sourcing Their AI Systems. Available online: <http://theconversation.com/why-big-tech-companies-are-open-sourcing-their-ai-systems-54437> (accessed on 7 October 2019).
4. Lopez, A.M., Jr.; Schulte, L.J.; Giguette, M.S. Climbing onto the shoulders of giants. *ACM SIGCSE Bull.* **2005**, *37*, 401–405. [[CrossRef](#)]
5. De Ribaupierre, H.; Jones, K.; Loizides, F.; Cherdantseva, Y. Towards Gender Equality in Software Engineering: The NSA Approach. In Proceedings of the 1st International Workshop on Gender Equality in Software Engineering, Gothenburg, Sweden, 28 May 2018; pp. 10–13.
6. Cheryan, S.; Master, A.; Meltzoff, A.N. Cultural stereotypes as gatekeepers: Increasing girls a interest in computer science and engineering by diversifying stereotypes. *Front. Psychol.* **2015**, *6*. [[CrossRef](#)] [[PubMed](#)]
7. Qiu, H.S.; Nolte, A.; Brown, A.; Serebrenik, A.; Vasilescu, B. Going farther together: The impact of social capital on sustained participation in open source. In Proceedings of the 41st International Conference on Software Engineering, Montreal, QC, Canada, 25–31 May 2019.
8. Durruthy, R. LinkedIn’s 2018 Workforce Diversity Report. Available online: <https://careers.linkedin.com/diversity-and-inclusion/workforce-diversity-report> (accessed on 7 October 2019).
9. Brown, D.; Parker, M. Google Diversity Annual Report 2019. Available online: <https://diversity.google/annual-report/> (accessed on 7 October 2019).
10. UNESCO Priority Gender Equality Action Plan 2014–2021. Available online: <https://unesdoc.unesco.org/ark:/48223/pf0000227222> (accessed on 7 October 2019).
11. Vasilescu, B.; Posnett, D.; Ray, B.; van den Brand, M.G.; Serebrenik, A.; Devanbu, P.; Filkov, V. Gender and Tenure Diversity in GitHub Teams. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; pp. 3789–3798. [[CrossRef](#)]
12. Mendez, C.; Sarma, A.; Burnett, M. Gender in Open Source Software: What the Tools Tell. In Proceedings of the 1st International Workshop on Gender Equality in Software Engineering, Gothenburg, Sweden, 28 May 2018; pp. 21–24. [[CrossRef](#)]
13. Izquierdo, D.; Huesman, N.; Serebrenik, A.; Robles, G. OpenStack Gender Diversity Report. *IEEE Softw.* **2019**, *36*, 28–33. [[CrossRef](#)]
14. Bonaccorsi, A.; Giannangeli, S.; Rossi, C. Entry strategies under competing standards: Hybrid business models in the open source software industry. *Manag. Sci.* **2006**, *52*. [[CrossRef](#)]
15. Rheingans, P.; Deramo, E.; Diaz-Espinoza, C.; Ireland, D. A Model for Increasing Gender Diversity in Technology. In Proceedings of the 49th ACM Technical Symposium on Computer Science Education, Baltimore, MD, USA, 21–24 February 2018. [[CrossRef](#)]
16. Botella, C.; Rueda, S.; López-Iñesta, E.; Marzal, P. Gender Diversity in STEM Disciplines: A Multiple Factor Problem. *Entropy* **2019**, *21*, 30. [[CrossRef](#)]
17. Olmedo-Torre, N.; Carracedo, F.S.; Ballesteros, M.N.S.; Lopez, D.; Perez-Poch, A.; Lopez-Beltran, M. Do Female Motives for Enrolling Vary According to STEM Profile? *IEEE Trans. Educ.* **2018**, *61*, 289–297. [[CrossRef](#)]
18. Vachovsky, M.E.; Wu, G.; Chaturapruek, S.; Russakovsky, O.; Sommer, R.; Fei-Fei, L. Toward More Gender Diversity in CS through an Artificial Intelligence Summer Program for High School Girls. In Proceedings of the 47th ACM Technical Symposium on Computing Science Education, Memphis, TN, USA, 2–5 March 2016. [[CrossRef](#)]
19. Geiger, R.S. Summary Analysis of the 2017 GitHub Open Source Survey. *arXiv* **2017**, arXiv:1706.02777.
20. Terrell, J.; Kofink, A.; Middleton, J.; Rainear, C.; Murphy-Hill, E.R.; Parnin, C.; Stallings, J. Gender differences and bias in open source: Pull request acceptance of women versus men. *PeerJ Comput. Sci.* **2017**, *3*, e111. [[CrossRef](#)]

21. Ford, D.; Smith, J.; Guo, P.J.; Parnin, C. Paradise unplugged: Identifying barriers for female participation on stack overflow. In Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering, Seattle, WA, USA, 13–18 November 2016; pp. 846–857.
22. Ford, D.; Harkins, A.; Parnin, C. Someone like me: How does peer parity influence participation of women on stack overflow? In Proceedings of the 2017 IEEE Symposium on Visual Languages and Human-Centric Computing, Raleigh, NC, USA, 11–14 October 2017; pp. 239–243.
23. May, A.; Wachs, J.; Hannak, A. Gender Differences in Participation and Reward on Stack Overflow. *Empir. Softw. Eng.* **2019**, *24*, 1997–2019. [[CrossRef](#)]
24. Kitchenham, B.A.; Dyba, T.; Jorgensen, M. Evidence-based software engineering. In Proceedings of the 26th International Conference on Software Engineering, Edinburgh, UK, 28 May 2004; pp. 273–281.
25. Felizardo, K.R.; Nakagawa, E.Y.; Fabbri, S.C.P.F.; Ferrari, F.C. *Revisão Sistemática da Literatura em Engenharia de Software: Teoria e Prática*; Elsevier: Rio de Janeiro, Brasil, 2017.
26. Brereton, P.; Kitchenham, B.A.; Budgen, D.; Turner, M.; Khalil, M. Lessons from applying the systematic literature review process within the software engineering domain. *J. Syst. Softw.* **2007**, *80*, 571–583. [[CrossRef](#)]
27. Biolchini, J.; Mian, P.G.; Natali, A.C.C.; Travassos, G.H. *Systematic Review in Software Engineering*; Technical Report ES; System Engineering and Computer Science Department COPPE/UFRJ: College Park, MD, USA, 2005.
28. Fabbri, S.; Silva, C.; Hernandez, E.; Octaviano, F.; Di Thommazo, A.; Belgamo, A. Improvements in the StArt tool to better support the systematic review process. In Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering, Limerick, Ireland, 1–3 June 2016; p. 21.
29. Silva, F.S.; Soares, F.S.F.; Peres, A.L.; de Azevedo, I.M.; Vasconcelos, A.P.L.; Kamei, F.K.; de Lemos Meira, S.R. Using CMMI together with agile software development: A systematic review. *Inf. Softw. Technol.* **2015**, *58*, 20–43. [[CrossRef](#)]
30. Petersen, K.; Vakkalanka, S.; Kuzniarz, L. Guidelines for conducting systematic mapping studies in software engineering: An update. *Inf. Softw. Technol.* **2015**, *64*, 1–18. [[CrossRef](#)]
31. Mani, A.; Mukherjee, R. A study of FOSS 2013 survey data using clustering techniques. In Proceedings of the 2016 IEEE International WIE Conference on Electrical and Computer Engineering, Pune, India, 19–21 December 2016; pp. 118–121. [[CrossRef](#)]
32. Tourani, P.; Adams, B.; Serebrenik, A. Code of conduct in open source projects. In Proceedings of the 2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering, Klagenfurt, Austria, 20–24 February 2017; pp. 24–33. [[CrossRef](#)]
33. Kuechler, V.; Gilbertson, C.; Jensen, C. Gender Differences in Early Free and Open Source Software Joining Process. In Proceedings of the IFIP International Conference on Open Source Systems, Hammamet, Tunisia, 10–13 September 2012; pp. 78–93.
34. Vedres, B.; Vasarhelyi, O. Gendered behavior as a disadvantage in open source software development. *EPJ Data Sci.* **2019**, *8*, 25. [[CrossRef](#)]
35. Moon, E. Gendered Patterns of Politeness in Free/Libre Open Source Software Development. In Proceedings of the 2013 46th Hawaii International Conference on System Sciences, Wailea, HI, USA, 7–10 January 2013; pp. 3168–3177.
36. Qiu, Y.; Stewart, K.J.; Bartol, K.M. Joining and Socialization in Open Source Women’s Groups: An Exploratory Study of KDE-Women. In Proceedings of the IFIP International Conference on Open Source Systems, Notre Dame, IN, USA, 30 May–2 June 2010; pp. 239–251.
37. Parra, E.; Haiduc, S.; James, R. Making a Difference: An Overview of Humanitarian Free Open Source Systems. In Proceedings of the 2016 IEEE/ACM 38th International Conference on Software Engineering Companion, Austin, TX, USA, 14–22 May 2016; pp. 731–733.
38. Davidson, J.L.; Naik, R.; Mannan, U.A.; Azarbakht, A.; Jensen, C. On older adults in free/open source software: Reflections of contributors and community leaders. In Proceedings of the 2014 IEEE Symposium on Visual Languages and Human-Centric Computing, Melbourne, VIC, Australia, 28 July–1 August 2014; pp. 93–100. [[CrossRef](#)]

39. Mendez, C.; Padala, H.S.; Steine-Hanson, Z.; Hildebrand, C.; Horvath, A.; Hill, C.; Simpson, L.; Patil, N.; Sarma, A.; Burnett, M. Open Source Barriers to Entry, Revisited: A Sociotechnical Perspective. In Proceedings of the 2018 IEEE/ACM 40th International Conference on Software Engineering, Gothenburg, Sweden, 27 May–3 June 2018; pp. 1004–1015. [CrossRef]
40. Nafus, D. ‘Patches don’t have gender’: What is not open in open source software. *New Media Soc.* **2012**, *14*, 669–683. [CrossRef]
41. Forte, A.; Antin, J.; Bardzell, S.; Honeywell, L.; Riedl, J.; Stierch, S. Some of All Human Knowledge: Gender and Participation in Peer Production. In Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion, Seattle, WA, USA, 11–15 February 2012; pp. 33–36. [CrossRef]
42. Mahmod, M.; Yusof, S.A.M.; Dahalin, Z.M. Women contributions to open source software innovation: A social constructivist perspective. In Proceedings of the 2010 International Symposium on Information Technology, Kuala Lumpur, Malaysia, 15–17 June 2010; pp. 1433–1438. [CrossRef]
43. Lee, A.; Carver, J.C. FLOSS participants’ perceptions about gender and inclusiveness: A survey. In Proceedings of the 41st International Conference on Software Engineering, Montréal, QC, Canada, 25–31 May 2019; pp. 677–687.
44. Cheryan, S.; Ziegler, S.A.; Montoya, A.K.; Jiang, L. Why are some STEM fields more gender balanced than others? *Psychol. Bull.* **2017**, *143*, 1–35. [CrossRef] [PubMed]
45. Robson, N. Diversity and Decorum in Open Source Communities. In Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, Lake Buena Vista, FL, USA, 4–9 November 2018; pp. 986–987. [CrossRef]
46. Silva, J.D.O.; Wiese, I.S.; German, D.M.; Steinmacher, I.F.; Gerosa, M.A. How Long and How Much: What to Expect from Summer of Code Participants? In Proceedings of the 2017 IEEE International Conference on Software Maintenance and Evolution, Shanghai, China, 17–22 September 2017. [CrossRef]
47. Robles, G.; Arjona Reina, L.; Serebrenik, A.; Vasilescu, B.; González-Barahona, J.M. FLOSS 2013: A Survey Dataset About Free Software Contributors: Challenges for Curating, Sharing, and Combining. In Proceedings of the 11th Working Conference on Mining Software Repositories, Hyderabad, India, 31 May–1 June 2014; pp. 396–399. [CrossRef]
48. Google Summer of Code 2019. Available online: <https://summerofcode.withgoogle.com/> (accessed on 7 October 2019).
49. Stack Overflows annual Developer Survey 2019. Available online: <https://insights.stackoverflow.com/survey/2019> (accessed on 7 October 2019).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).