Gender Diversity and Women in Software Teams: How Do They Affect Community Smells?

Authors: Gemma Catolino, Fabio Palomba, Damian A. Tamburri, Alexander Serebrenik, Filomena Ferrucci

Abstract

The study aimed to assess the impact of gender diversity and women's participation on team performance in software development projects. The study used a mixed-methods research design, combining an online survey and interviews with project managers to collect data from software development communities worldwide. The study found that gender diversity and women's participation can have a positive impact on team performance through better communication, cooperation, and fewer community smells.

Population Targeted:

The population targeted in the study was software development communities, including both open-source and closed-source projects. The unique traits or characteristics of this population are that they are composed of individuals with diverse technical backgrounds, skills, and experiences who work collaboratively to develop software products. These communities often have a high degree of pluralism, meaning that they consist of individuals from different cultural, social, and economic backgrounds who bring their unique perspectives and ideas to the development process. The study focused on gender diversity and women's participation within these communities as a way to explore how pluralism can impact team performance in software development projects.

Goal of Study, Research Questions, and Hypothesis:

- The goal of the study was to assess the impact of gender diversity and women's participation on team performance in software development projects.

- The research questions asked were:
  1. To what extent does the presence of women within teams influence the number of community smells?
  2. How does the number of community smells differ in teams without women and teams with women?

- The hypothesis:
  The presence of women within the team improves communication, co-operation, collaboration, thus reducing the number community smells.

Research Methods Used:

- The study used a mixed-methods research design, combining an online survey and interviews with project managers to collect data from software development communities worldwide. The online survey was used to collect quantitative data on gender diversity, women's participation, and community smells within software development teams. The interviews with project managers were used to gather qualitative data on their experiences with gender diversity and women's participation in software development teams.

Outcomes:

- The study found that gender diversity and women's participation can have a positive impact on team performance through better communication, cooperation, and fewer community smells within software development teams. The study also found that there is a need for more research on how organizational policies and practices can promote diversity and inclusion within software development teams.

Literature background

- The literature review highlights that gender diversity has been shown to have a positive impact on organizational performance and innovation. However, there is limited research on the specific role of gender diversity in software development teams. The review also notes that women are underrepresented in software engineering, which can lead to a lack of diverse perspectives and hinder innovation.

- The authors cite several studies that have found a positive correlation between gender diversity and team performance, as well as studies that have identified communication and collaboration as key factors for successful software development teams. The literature review also discusses some challenges associated with achieving gender balance in software engineering, such as unconscious bias and stereotypes.

Key aspects of research design

1. Data Collection: The study collected data from 20 different software development communities, each with at least 10 members. The data was collected through online surveys and interviews with project managers.

2. Variables: The study measured gender diversity and women's participation in each community, as well as the number of "community smells" present. Community smells are defined as indicators of poor software development practices or code quality issues.

3. Analysis: The study used statistical analysis to examine the relationship between gender diversity, women's participation, and community smells. The analysis included regression models and correlation tests.

4. Mitigating Threats to Validity: The study identified several potential threats to validity, such as selection bias and social desirability bias, and took steps to mitigate them. For example, the study used random sampling to select communities for inclusion in the study and used anonymous surveys to reduce social desirability bias. Overall, the research design aimed to provide a rigorous empirical analysis of the relationship between gender diversity, women's participation, and software development team performance while addressing potential threats to validity.

Key Aspects of Research Methods:

- The study used a mixed-methods research design, combining an online survey and interviews with project managers to collect data from software development communities. The study used statistical analysis to examine the relationship between gender diversity, women's participation, and community smells. The online survey was used to collect quantitative data on gender diversity, women's participation, and community smells within software development teams. The survey questions were designed to measure the extent of gender diversity within teams, the level of women's participation in software development projects, and the prevalence of community smells within teams. The interviews with project managers were used to gather qualitative data on their experiences with gender diversity and women's participation in software development teams.

Community smells

- **Organizational Silo Effect:** This smell appears when a software community has siloed areas that do not communicate with each other, except through one or two members. This reflects rigid thinking and lack of community-wide communication. The presence of women is expected to break down the barriers forming across community silos, thus reducing the amount of organizational silo effects.

- **Black Cloud Effect:** This smell appears when there is an excessive information overload due to the lack of structured communication. The Black Cloud effect corresponds to the manifested presence of overwhelming quantities of asynchronous and synchronous data exchanges across a community. The presence of more diversity is expected to mediate positively on the presence of black clouds, and it might have an effect with respect to the tenor of communications.

Community smells

- **Lone Wolf Effect:** This smell appears when there are unsanctioned or defiant contributors who carry out their work irrespective or regardless of their peers. The presence of women is expected to have a positive influence and reduce the amount of reported "lone wolves." However, this mitigating effect is expected to manifest only if women have had a direct involvement in the process of designing and developing the involved software.

- **Radio Silence Effect:** This is an instance of the "unique boundary spanner" problem from social-networks analysis: one member interposes herself into every formal interaction across two or more sub-communities with little or no flexibility to introduce other parallel channels. Promotion-rates for women to cover boundary-spanning roles have previously been discovered and established, and thus their role is expected to mediate for the presence of such a community smell.

Community smells

It is worth noting that the considered smells can be grouped into two high-level categories, i.e., structural- and communication-based. While *Organizational Silo* and *Black- cloud* affect the overall community structure, *Lone Wolf* and *Radio Silence* are concerned with the way the community members communicate with each other. Thus, the selection of those smells allows us to understand what is the role of gender diversity and the presence of women on problems having two different levels of granularity.

Validation

The researchers evaluated the effectiveness of the detection techniques by running CODEFACE4SMELLS on 60 open-source projects and surveying the original developers to determine if the identified community smells reflected real community-related problems. The researchers found that the tool's recommendations highlighted real community-related problems, and the effectiveness of the operationalizations relied on the proven effectiveness of the approach by Joblin et al. [21] building upon the "Order Statistics Local Optimization Method" (OSLOM) [22] featured inside CODEFACE.

Research Question-1

- They first completed the detection of community smells for each of the time windows considered using CODEFACE4SMELLS. Then, they computed statistical metrics such as minimum, maximum, mean, median, first, and third quartile of the distribution of community smells in the two groups. The aim was to investigate whether non-gender-diverse teams present a higher number of community smells with respect to gender-diverse ones. They used the Mann-Whitney test to statistically assess the differences observed between the two groups and measure the effect size with Cliff's Delta (d). The authors followed well-established guidelines to interpret the results based on the magnitude of the effect size.

Research Question-2

- The statistical model relates the community metrics present in the dataset to the detected smells. The independent variables considered in the model are the number of women in a team and the Blau-Index. Control variables that were considered include the number of committers, number of commits, team size, turnover, project age, and tenure diversity. The response variable was the number of community smells identified in each of the time windows of the considered projects. Control variables such as socio-technical congruence, number of committers, and tenure median were also considered in the study. The study aimed to analyze the impact of women participation and gender diversity on the presence of community smells.

Research Question-1 results

- The study found that teams without women had a higher number of community smells compared to teams with women. The mean for community smells in teams without women was 7, while it was 3 in teams with women. The difference in the distributions was statistically significant with a large effect size, and the results were controlled for team size. This indicates that there may be factors within gender-diverse teams that influence the number of community smells, motivating further analysis in RQ2 to identify these factors.

Research Question-2 results

- They compared their model with two baseline statistical models to better analyze and interpret their findings. They found that the presence of women was a significant mediator for the occurrence of the Radio Silence smell, while for Organizational Silo and Lone Wolf, the presence of women played a mild mediating role at best. Team size and socio-technical congruence were found to be more relevant factors for Lone Wolf. For Black Cloud, the Blau-index, which denotes diversity, was found to have the strongest estimate and significantly reduced the number of instances of this smell type. Overall, the authors conclude that the presence of women can only partially mediate the occurrence of community smells, and having a diverse team significantly helps in reducing the number of Black Cloud instances.

Threats to validity

- Overall, the study has several threats to validity that should be taken into consideration when interpreting the results. These threats include the possible imprecision of variables computed from the dataset, the limitations of the CODEFACE4SMELLS tool used to detect community smells, and the binary classification of gender. Additionally, there may be threats to conclusion validity related to the statistical methods employed, such as the possibility of multi-collinearity, the presence of outliers, and the use of ANOVA to interpret the results. Finally, there are threats to external validity, as the study was limited to a specific set of software systems and may not be generalizable to other contexts. Future research should aim to address these limitations and expand the scope of the study.

My evaluation

- Strengths:

  The mixed-methods research design used in the study allowed for a comprehensive understanding of how gender diversity and women's participation impact team performance in software development projects. The use of both quantitative and qualitative methods provided a more complete picture of the research questions, allowing for triangulation of data and increased validity. The online survey used in the study had a large sample size, which increases the generalizability of the findings. The survey questions were designed to measure specific variables related to gender diversity, women's participation, and community smells within software development teams, which increases the reliability of the data collected. The interviews with project managers provided valuable qualitative data on their experiences with gender diversity and women's participation in software development teams. The interviews allowed for an in-depth exploration of how gender diversity impacts team performance and innovation.

My evaluation

- Weaknesses:

  One potential weakness of the study is that it relied on self-reported data from participants, which may be subject to bias or inaccuracies. For example, participants may have over-reported their level of gender diversity or under-reported the prevalence of community smells within their teams. Another potential weakness is that the study did not control for other factors that may impact team performance, such as team size or project complexity. This limits the ability to draw causal conclusions about the relationship between gender diversity, women's participation, and team performance. Finally, while the online survey had a large sample size, it may not be representative of all software development communities worldwide. The study relied on convenience sampling methods (e.g., social media platforms and email listservs), which may have introduced selection bias into the sample.

Major Findings

1. Gender Diversity and Women's Participation: The study found that software development communities with higher levels of gender diversity and women's participation had fewer community smells than those with lower levels.

2. Communication and Cooperation: The study found that gender diversity and women's participation were positively correlated with better communication and cooperation within software development teams.

3. Community Smells: The study found that community smells were more prevalent in software development communities without women than in those with women.

4. Project Managers' Perceptions: The study found that project managers who had experience working with mixed-gender teams generally had positive perceptions of gender diversity and women's participation in software development teams.

Overall, the study suggests that gender diversity and women's participation can have a positive impact on software development team performance, specifically in terms of reducing community smells and improving communication and cooperation within teams.

Discussion points/implications

1. Importance of Gender Diversity: The study highlights the importance of gender diversity in software development teams, as it can lead to better team performance and innovation.

2. Addressing Gender Imbalance: The study suggests that addressing gender imbalance in software engineering is crucial for achieving better team performance. This can be done by promoting diversity and inclusion initiatives, reducing unconscious bias, and increasing awareness of the benefits of gender diversity.

3. Communication and Cooperation: The study emphasizes the importance of communication and cooperation within software development teams for achieving better team performance. Project managers should focus on creating an environment that encourages open communication and collaboration among team members.

4. Future Research: The study suggests several areas for future research, such as exploring the impact of other types of diversity (e.g., cultural diversity) on software development team performance, investigating the role of leadership in promoting gender diversity, and examining how gender diversity affects different stages of software development projects. Overall, the study has important implications for promoting gender diversity in software engineering and improving team performance through better communication and cooperation.
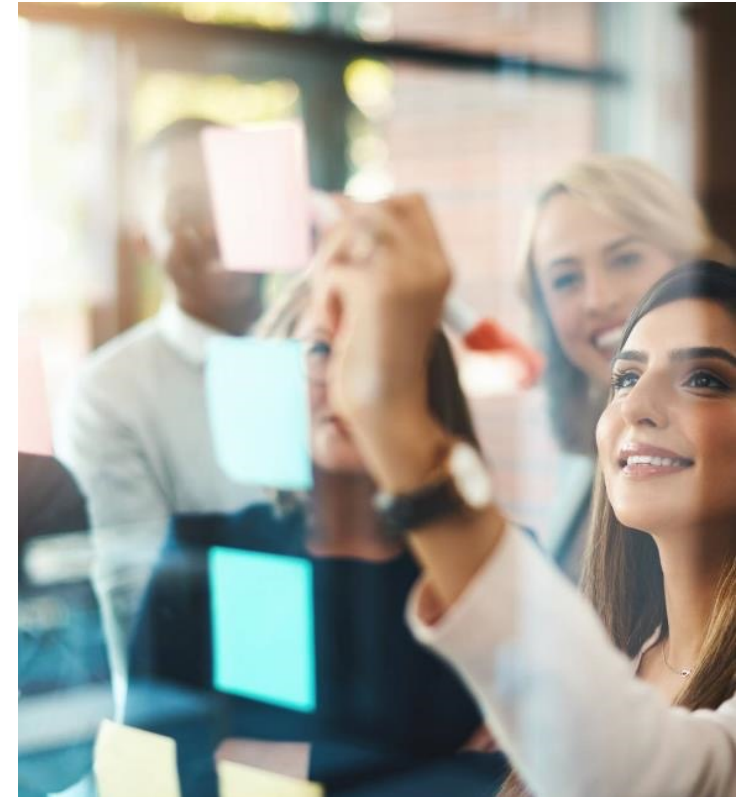
Thoughts/Reflections

- Overall, the study provides valuable insights into the relationship between gender diversity, women's participation, and team performance in software development projects. The mixed-methods research design used in the study allowed for a comprehensive understanding of the research questions, and the large sample size of the online survey increases the generalizability of the findings. However, there are some potential weaknesses in the study that should be considered when interpreting the results. For example, relying on self-reported data from participants may introduce bias or inaccuracies into the data. Additionally, not controlling for other factors that may impact team performance limits the ability to draw causal conclusions about the relationship between gender diversity, women's participation, and team performance.
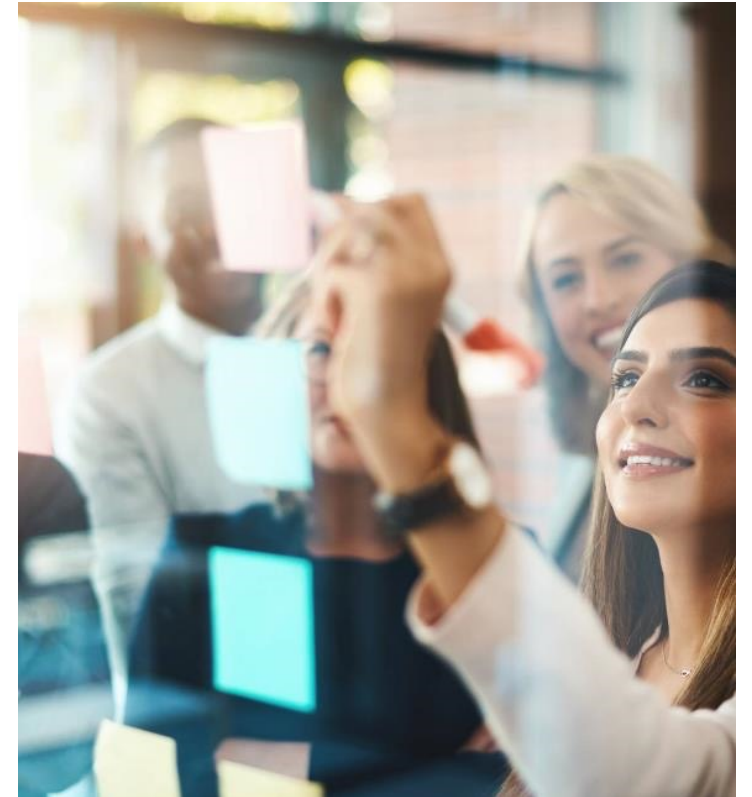
Thoughts/Reflections

- Alternative approaches to this study could include using experimental methods to manipulate levels of gender diversity or women's participation within software development teams and measuring their impact on team performance. This would allow for stronger causal conclusions about the relationship between these variables. To extend or redo this work, future studies could explore other types of diversity (e.g., cultural diversity) and their impact on team performance in software development projects. Additionally, future studies could examine how organizational policies and practices can promote diversity and inclusion within software development teams. Finally, future studies could explore how gender diversity and women's participation impact specific aspects of team performance (e.g., innovation or productivity).

THE END