

# Investigating the Impact of Interpersonal Challenges on Feeling Welcome in OSS

Bianca Trinkenreich\*, Zixuan Feng<sup>†</sup>, Rudrajit Choudhuri<sup>‡</sup>, Marco Gerosa<sup>‡</sup>, Anita Sarma<sup>†</sup>, Igor Steinmacher<sup>‡</sup>

\*Colorado State University, United States, {bianca.trinkenreich}@colostate.edu

<sup>†</sup>Oregon State University, United States, {fengzi, choudhru, anita.sarma}@oregonstate.edu

<sup>‡</sup>Northern Arizona University, United States, {marco.gerosa,igor.steinmacher}@nau.edu

**Abstract**—The sustainability of open source software (OSS) projects hinges on contributor retention. Interpersonal challenges can inhibit a feeling of welcomeness among contributors, particularly from underrepresented groups, which impacts their decision to continue with the project. How much this impact is, varies among individuals, underlining the importance of a thorough understanding of their effects. Here, we investigate the effects of interpersonal challenges on the sense of welcomeness among diverse populations within OSS, through the diversity lenses of gender, race, and (dis)ability. We analyzed the large-scale Linux Foundation Diversity and Inclusion survey (n = 706) to model a theoretical framework linking interpersonal challenges with the sense of welcomeness through Structural Equation Models Partial Least Squares (PLS-SEM). We then examine the model to identify the impact of these challenges on different demographics through Multi-Group Analysis (MGA). Finally, we conducted a regression analysis to investigate how differently people from different demographics experience different types of interpersonal challenges. Our findings confirm the negative association between interpersonal challenges and the feeling of welcomeness in OSS, with this relationship being more pronounced among gender minorities and people with disabilities. We found that different challenges have unique impacts on how people feel welcomed, with variations across gender, race, and disability groups. We also provide evidence that people from gender minorities and with disabilities are more likely to experience interpersonal challenges than their counterparts, especially when we analyze stalking, sexual harassment, and doxxing. Our insights benefit OSS communities, informing potential strategies to improve the landscape of interpersonal relationships, ultimately fostering more inclusive and welcoming communities.

**Index Terms**—challenges, OSS, diversity, open source, OSS, toxicity, inclusivity, belonging, welcomeness

## I. INTRODUCTION

A core principle of Open Source Software (OSS) is its openness—anyone from anywhere in the world can contribute, irrespective of their background. However, despite this apparent openness, OSS projects have a severe lack of diversity. For example, data reveals that less than 10% of the contributors are women [1], and rates are also low for non-whites (16.6%) [2] and people who have disabilities (17%) [3]. This lack of diversity is not just a social justice problem, it also undermines the innovation and sustainability of OSS projects [4, 5].

Literature has also extensively explored the multiple facets of challenges faced by OSS contributors [6–9], including barriers faced by newcomers [6], one-time contributors [10], and mentors [11]. Prior work has investigated how a Sense of Virtual Community (SVC) in OSS is created by intrinsic factors

such as membership, identity, and attachment to a group and affects their sense of belonging [12]. However, [12] does not tie this feeling of SVC to the challenges that individuals face, which are the antecedents of SVC. More specifically, studies have identified interpersonal challenges as a major obstacle to contributing to OSS [13], especially for those who are a minority in OSS [1, 14, 15].

What is missing, however, is an understanding of how these interpersonal challenges hinder the contributors' feeling of welcomeness in the project, which in turn impacts their decision to continue contributing. This oversight is not trivial; the feeling of welcomeness is key to sustained engagement, particularly in environments built on open collaboration principles [16]. This brings us to our first research question: **RQ1:** How are interpersonal challenges associated with the feeling of welcomeness in OSS?

Research has also shown that individuals in the minority in OSS often face a higher frequency of interpersonal challenges such as toxic communication or gender stereotyping, which in turn impacts their turnover [1, 11, 17]. However, the extent to which interpersonal challenges affect the feeling of welcomeness among minority communities is yet to be determined. Additionally, our worldview is an amalgamation of our individual backgrounds and lived experiences. Consequently, various interpersonal challenges may exert disparate influences on individuals belonging to distinct demographic groups. Thus far, no research has quantified the differential effects of various challenges on feeling welcome nor on the interpersonal challenges across demographic dimensions. This leads us to our other two research questions: **RQ2:** How does the association between interpersonal challenges and feeling welcome differ for those who are in gender, race, or disability minorities? **RQ3:** How differently do those in gender, race, or disability minority face interpersonal challenges?

The current absence of a comprehensive understanding that delineates the impact of challenges on the feeling of welcomeness across various population segments implies that endeavors by OSS communities to retain a diverse pool of contributors are likely to fail.

Our study closes this gap through a large-scale (n=706), mixed-methods analysis of Linux Foundation survey data [18]. Specifically, we build a theoretical model using Partial-Least Squares Structural Equation Modeling (PLS-SEM) to determine and delineate the different types of interpersonal challenges,

their extent, and their impact on feeling welcome (RQ1). We then model the differential impact of the various interpersonal challenges on demographic dimensions of gender, race, and disabilities through Multi-Group Analysis (RQ2). Following, we modeled individual logistic regressions to evaluate the differences in experiencing interpersonal challenges across demographic dimensions of gender, race, and disabilities (RQ3).

Our study follows a concurrent embedded strategy [19] and uses qualitative analysis of the open-ended survey responses about reasons to feel (or not feel) welcome to understand the participants' specific experiences that contributed to their feeling (or lack) of welcomeness in the community. This serves as a form of data triangulation to provide rigor to our findings.

Our theoretical model empirically shows that experiencing interpersonal challenges is negatively associated with feeling welcome in OSS. Moreover, it highlights the nuanced nature of how different challenges impact different demographics differently. For example, in our data, while race minorities felt less welcome, interpersonal challenges had a strong association with not feeling welcome for gender minorities and people with disabilities, but not for race.

By understanding how specific challenges impact diverse contributors' feelings of welcomeness, our study not only advances the academic literature on the field of behavioral software engineering but also holds practical implications. The results are paramount to inform OSS communities on the specific aspects to focus on to foster diversity in OSS, which in turn can result in innovation, and drive the sustainability of OSS projects. Our contributions offer valuable insights that may go beyond OSS, informing broader software industry efforts on creating inclusive and equitable spaces.

## II. RELATED WORK

**Diversity and Inclusion in OSS.** Diversity and inclusion in OSS have received attention in recent years [1, 4, 5, 12, 17, 20–23]. Most of the literature related to minorities and underrepresentation in OSS is focused on women [1, 4, 5, 17, 22, 24]. A literature survey by Trinkenreich et al. [1] identified that women are poorly represented and aggregates eight significant challenges women face when contributing to OSS. By analyzing repositories, Vasilescu et al. [4] showed that projects with higher gender diversity within their development teams tend to experience more effective collaboration and productivity. Similarly, a survey by Lee and Carver [17] highlighted notable perceptions of gender bias and inclusivity issues within the OSS community. Participants acknowledged barriers to women's participation and emphasized the need for better practices to foster inclusivity.

However, fewer studies have focused on other underrepresented populations in OSS besides gender. Previous research claimed for more studies about racial diversity in software engineering [14]. Race minority was evidenced as underrepresented in OSS [25, 26]. In this context, Nadri et al. [15] found that perceptibly Black and Asian/Pacific developers faced higher contribution rejection rates than those from perceptibly White developers. A literature review by Rodríguez-Pérez et al.

[14] indicates limited knowledge about the impact of race, nationality, and disability in OSS. Despite this, respondents in a recent survey reported that their disability influences their emotions and perceptions of others' feelings [27].

**Interpersonal challenges in OSS.** "Open source has a reputation for being aggressive" [28]. Toxicity and incivility are prevalent issues in OSS discussions and code reviews, adversely affecting contributor engagement, collaboration, and overall community health Miller et al. [29], Ferreira et al. [30]. Contributors who are underrepresented in OSS are experiencing even worse collaboration environments. Harassing verbal comments, sexual jokes, insults, and bullying are not uncommon [31, 32]. Singh and Brandon [33] found that less than 5% of online communities are "safe" to women contributors, free of sexism and discrimination. However, from a quantitative analysis of 355 OSS software package websites, only 10% of them have a Code of Conduct or relevant rules [34]. Such prevalent interpersonal challenges are creating barriers that make contributors feel unwelcome in OSS, especially for contributors who are underrepresented.

**Welcomeness in OSS.** Feeling welcome antecedes belonging [35]. Gill [36] describes welcomeness as more than a mere acknowledgment of arrival or permission for entry. It embodies a deeper sense of emotional and relational connection between the welcomer and the newcomer, implying a willingness to engage and suggesting a deeper level of engagement beyond mere permission. It involves an exchange of emotions that transcend bureaucratic processes [36].

We introduce the concept of welcomeness in OSS, which focuses on conveying a genuine positive reception of newcomers as contributors to the project. While onboarding primarily addresses the procedural and technical requirements necessary to contribute, welcomeness emphasizes the interpersonal dynamics that foster an engaging atmosphere characterized by mutual respect and connection.

Welcomeness is a construct investigated in the context of the inclusion of minorities in other areas, like, A [37], Education [38, 39], Sports [38], International Relations [40], and Healthcare [41]. While there is no previous work specifically focusing on welcomeness in software engineering, it antecedes the sense of belonging [35]—which was the focus on recent studies in software engineering teams context [12, 42, 43].

The need for belonging is crucial in the OSS domain, as shown in studies of OSS contributors [12, 23]. More specifically, close to the present work, Trinkenreich et al. [12] showed that the sense of virtual community plays a critical role in the contributors' continued engagement and contributions. Trinkenreich's study highlights that factors such as shared experiences, support, and meaningful interaction within the community are vital in fostering a sense of belonging among contributors, thereby influencing their commitment to the OSS environment. Therefore, the sense of belonging emerges as a key factor in retaining personnel, which is vital for the sustainability of OSS projects.

While previous research extensively investigates the various aspects analyzed in this study, it does not connect aspects of

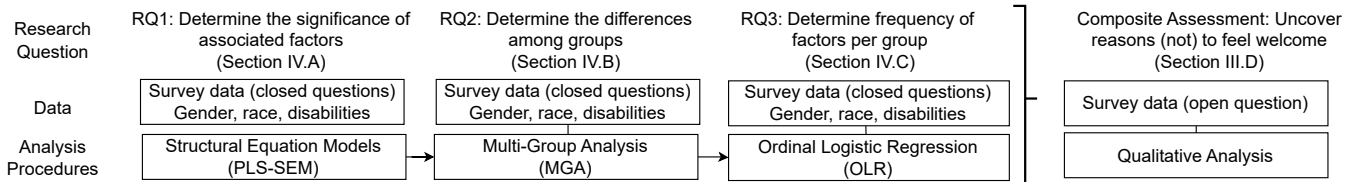


Fig. 1: Research design following a Concurrent Embedded Mixed Methods Strategy

belonging with the interpersonal challenges faced by contributors nor consider the feeling of welcomeness as a distinct construct. Our present work complements the existing literature by evidencing the interplay between interpersonal challenges and welcomeness—which anticipates belonging [35]—and their effect on gender, race, and disabilities minorities.

### III. RESEARCH DESIGN

Fig. 1 presents an overview of the research design. We followed a Concurrent Embedded Mixed Methods Strategy [19] with a dominant quantitative component.

Answering RQ1 involves using Structural Equation Models - Partial Least Squares to model how interpersonal challenges are associated with the feeling of welcomeness. In RQ2, we investigated the heterogeneity in the model through Multi-Group Analysis (MGA) and established whether the association of the variables identified in RQ1 varied according to gender, race or ethnicity, and abilities. RQ3 identifies the frequency of each demographic group experiencing each interpersonal challenge using Ordinal Logistic Regression. Finally, we qualitatively analyzed the open-ended survey responses about why contributors (do not) feel welcome in OSS as a form of data triangulation and conducted a follow-up study to collect the practitioners’ perspectives about our results.

This work is different from the model of sense of belonging proposed by Trinkenreich et al. [12], which investigated the association between intrinsic motivations and the Sense of Virtual Community (SVC). Here we investigate the feeling of welcomeness, which anticipates belonging [35], and its association with challenges, not motivations. Moreover, Trinkenreich et al. [12]’s study used the diversity lens of gender, while the present study also tackles the lens of race and disabilities.

#### A. Dataset

We used the data from the Linux Foundation Diversity and Inclusion survey [18]. The survey design was led by the Linux Foundation Research, in collaboration with 40 industrial partners and researchers from our community [18]—including an author of this paper, who provided ideas and feedback on the questionnaire. This collaboration between industry and academic researchers helped increase the instrument’s clarity, reliability, and reach. The questionnaire was available between July 15 and July 29, 2021, in multiple languages, but with the majority of responses (89%) in English.

The survey was advertised via social media and through the channels of the project partners (e.g., Linux Foundation, Amazon, CHAOSS Community, Comcast, Fujitsu, GitHub, GitLab, Hitachi, Huawei, Intel, NEC, Panasonic, Red Hat, Renesas, VMware, etc.). The survey was targeted at open source contributors. The compensation for answering the survey was

a 20% discount when registering to attend the Open Source Summit Conference. The survey received 2,350 responses.

Linux Foundation Research made the questionnaire and the quantitative data publicly available [44], and we obtained the qualitative data by contacting Linux Foundation Research.

The survey included a question about the feeling of welcomeness, a 5-point Likert scale item (from Strongly Disagree to Strongly Agree, with neutral). The questions about the frequency of experiencing challenges were asked as 4-point items: Never, Rarely, Occasionally, and Frequently. For the gender and race questions, respondents could mark as many options they identified with. The disabilities question could be answered as *yes* or *no*. Since our analysis techniques require complete responses, we removed responses with missing values, as they represented more than 10% of responses, and imputation methods could bias results for multi-group analyses used in RQ2 [45]. Given the large number of responses, the trade-off between introducing a potential bias versus increasing the usable sample size suggests a more conservative approach, so we did not use imputation methods. We removed the responses with blank values for: the feeling of welcomeness question (n=4); any interpersonal challenge items (n=104); gender (n=30); race (n=1503); and disabilities (n=2).

Post filtering, our dataset included 706 responses for the analyses. We present demographics in Table I.

TABLE I: Demographics of respondents (n=707)

Attribute	N	Percentage
Gender		
<b>Only Men</b>	<b>539</b>	<b>76.3%</b>
<b>Gender Minorities</b>	<b>167</b>	<b>23.7%</b>
Women	115	16.3%
Non-Binary or Third Gender	29	4.1%
More than one gender	23	3.3%
Disability		
<b>No</b>	<b>549</b>	<b>77.8%</b>
<b>Yes</b>	<b>157</b>	<b>22.2%</b>
Race		
<b>Only White</b>	<b>486</b>	<b>68.8%</b>
<b>Racial Minorities</b>	<b>220</b>	<b>31.2%</b>
Black	50	7.1%
Hispanic	43	6.1%
Asian	43	6.1%
Native or Indigenous	4	0.6%
Pacific Islander	2	0.3%
More than one race	78	11.0%

We added the subset of questions used in the study in the replication package [46], together with prepared datasets to run RQ1, RQ2, and RQ3, and the codebook used for qualitative analysis. For the sake of confidentiality between researchers and The Linux Foundation Research, the raw data from the open question could not be shared.

## B. Hypothesis Development

In the context of this work, **Feeling Welcomeness** refers to the self-perception of an individual feeling that their presence is desired, accepted, and valued by their OSS colleagues without acts of microaggressions [47, 48, 48]. Welcomeness is an inherently emotional and relational concept that entails more than simply being permitted to enter and includes inter-relational subjectivity [36]. Understanding the feelings of welcomeness in a given context plays an important role in identifying where barriers exist and how to break down these barriers to increase diversity in participation [38].

Whether an OSS community has a welcoming atmosphere is determined by the challenges that individuals face during the contribution process [11], affecting both the retention of existing and the attraction of new contributors. Interpersonal challenges are difficulties or conflicts that arise from interpersonal relationships (or lack thereof) in OSS projects [6, 17], and tend to have a more pronounced impact on demographics typically underrepresented in the tech industry [1, 11, 17]. These challenges can stem from differences in communication styles, cultural backgrounds, personal behaviors, and social dynamics, and can negatively impact collaboration, productivity, and the overall well-being of the individuals involved [6].

In the questionnaire, interpersonal challenges were defined based on (1) the unacceptable behaviors listed in the Linux Foundation's code of conduct [49] (unwelcoming language, stalking, sexual harassment), (2) issues gathered from the literature and seen by practitioners (lack of response to the rejection of contributions, stereotyping, doxxing), and (3) an overarching conflict-related question (conflict or tension with other contributors).

Given the potential relationship between interpersonal challenges and the feeling of welcomeness, we propose the following hypothesis:

**Hypothesis 1 (H1).** *Experiencing interpersonal challenges is negatively associated with feeling welcome.*

We used Partial Least Squares–Structural Equation Modeling (PLS-SEM) [50] to analyze the relationship between interpersonal challenges and feelings of welcomeness in OSS. SEM facilitates the simultaneous analysis of relationships among constructs that can be measured by one or more indicator variables. The main advantage of SEM is being able to measure complex model relationships while accounting for measurement error when using latent variables (e.g., Interpersonal Challenges).

In the following, we discuss the measurement model (i.e., operationalization of constructs) and analysis.

**Interpersonal challenge** is a theoretical concept that cannot be directly observed given its multiple facets. Therefore, this concept is represented as a latent variable. A latent variable cannot be directly measured but instead is measured through a set of indicators or manifest variables. The interpersonal challenges were measured in our dataset via seven 4-point Likert scale items representing different interpersonal challenges.

1. *"Doxxing"*, derived from *"document dropping,"* involves the public disclosure of private or personal information about an individual or organization, often with harmful intentions. Regarded as a detrimental and unethical action, doxxing can result in privacy breaches and, in some cases, physical harm to the affected parties [51].

2. *Sexual harassment* is a well-documented social issue that can occur offline or online with profound impacts on individuals and organizations. In OSS, conferences and meetups serve as platforms for in-person interaction and can expose minority attendees to inappropriate behavior. Sexual harassment can occur online also. For instance, when a newcomer who identifies as a woman seeks mentorship from a man, her intentions are often responded to as a dating opportunity [52].

3. *Stalking* refers to actions meant to intimidate or instill fear [53, 54], which happens in OSS as a form of toxicity [55].

4. *Stereotyping* represents discrimination based on perceived demographic characteristics [56], usually referring to a form of negative and fixed impression that can happen implicitly or explicitly and relate to a socially shared set of beliefs about traits of members of a social category [56].

5. *Unwelcoming Language* is defined as written or spoken communication that induces feelings of exclusion, including instances of profanity, racist jokes, and name-calling. Those who are in minority are more likely to experience unwelcoming language [1, 32, 52, 57, 58].

6. *Conflict or tension with other contributors* can manifest in diverse ways and may stem from variations in technical approaches, communication styles, or personal interactions. For instance, tension may arise due to misunderstandings, misinterpretations, or the use of ineffective communication channels [59]. Technical debates, such as differing opinions, design philosophies, or coding practices [60], can also be sources of tension among contributors.

7. *Lack of response to or rejection of contributions or questions* refers to situations where contributors do not receive acknowledgment or approval for their work or when questions are met with indifference, silence, or dismissal. Delayed or absent responses lead to contributor attrition in OSS [29, 61, 62].

## C. RQ1 Analysis

To answer RQ1, we quantitatively analyzed the responses to the questions about interpersonal challenges and the feeling of welcomeness using Structural Equation Models–Partial Least Squares (PLS-SEM), utilizing SmartPLS version 4. The first step in evaluating a structural equation model is to assess the soundness of the measurement of the latent variable—this is referred to as evaluating the ‘measurement model’ [63]. The second step is to evaluate the theoretical model, which includes the evaluation of the hypothesis. The measurement model is evaluated through tests of convergent validity, internal consistency reliability, discriminant validity, and collinearity. The theoretical model is evaluated according to the significance of the path coefficient and model fit, which are obtained through a ‘bootstrapping’ procedure: it draws a large number (e.g., 5,000) of random ‘subsamples’ of the same size as the original

sample (using replacement). The model is estimated for each subsample, generating a sampling distribution used to determine a standard error [63] that can subsequently be used to make statistical inferences. The model fit is then evaluated using Standardized Root Mean Square Residual (SRMR).

#### D. RQ2 Analysis

To answer RQ2, we evaluated the impact of interpersonal challenges (as a single construct) on the feeling of welcomeness across demographic groups. To define the demographic groups of interest, we followed the core categories identified by Albusays et al. [64]. Gender and race are very relevant perspectives in OSS, with many communities implementing initiatives specifically focused on these populations, whereas disability remains underexplored, likely due to data scarcity [65]. By focusing on these categories, we sought relevance, maintained a clear focus on the paper, and avoided data dredging (i.e., indiscriminately including all data).

We employed Multi-Group Analysis (MGA) to understand how the model varies within distinct demographic groups. We followed Hair Jr et al. [66] proposed structure that includes testing for invariance and comparing groups to analyze results.

*Testing for Measurement Invariance of Composite Models (MICOM)* is a mechanism to assess whether or not the loadings of the items representing the latent variables differ significantly across different groups. In other words, we want to establish whether any differences can be attributed to the theoretical constructs and not how we measured those constructs [66]. The literature recommends conducting MICOM before MGA to confirm its appropriateness [66]. If we find differences in the measurement model for the two groups, then we cannot attribute any differences in the model to the theoretical constructs.

Comparing group-specific model relationships for significant differences using a multi-group analysis requires establishing configural and compositional invariance [66, 67]. Configural invariance does not include a test and is a qualitative assessment to ensure all of the composites are equally defined (“configured”) for all of the groups, such as equivalent indicators per measurement model, equivalent treatment of the data, and equivalent algorithm settings or optimization criteria. Compositional invariance exists when the composite scores are the same across both groups; while small differences will naturally happen for different groups, we tested whether those differences are significantly different. We used MICOM procedure to examine the correlation between the composite scores of groups and assess that the correlation equals 1.

*Groups Comparison and Analysis* was done through bootstrapping multigroup analysis, and the statistical differences are assessed through parametric tests between path coefficients generated from different samples.

#### E. RQ3 Analysis

Results from RQ1 showed that interpersonal challenges are associated with less feeling of welcome, and RQ2 showed that this association is more prevalent for gender minorities and people with disabilities. However, much is not known about the frequency of challenges each demographic experiences.

To answer RQ3, we built seven individual logistic regression models, with each interpersonal challenge as the dependent variable for each model. The independent binary variables are gender minorities, race minorities, and disabilities (Table I), allowing us to assess which group will more likely encounter each interpersonal challenge. We used ordinal logistic regression since the dependent variables were from a 4-point Likert-scale response (never, rarely, occasionally, and frequently). The ordinal logistic regression is an extension of the logistic regression that can be used when the dependent variable is ordinal [68].

#### F. Composite Assessment

We qualitatively analyzed the open-ended responses to the question about the reasons for (not) feeling welcome by inductively applying open and axial coding. We built post-formed codes as the analysis progressed and associated them with respective parts of the answer to code the reasons according to the respondents’ perspectives. Reasons to not feel welcome were related to six out of the seven interpersonal challenges, and one technical challenge. We also found a positive reason to feel welcome. We used a negotiated agreement protocol and discussed the codes and categorization as a team until reaching a consensus about the themes, as cataloged in our codebook that is part of our replication package [46].

From the 706 answers we used for quantitative analysis, we received 588 answers for the open question of reasons to (not) feel welcome, and we identified respondents as P1 to P588.

#### G. Follow-up: Community Perceptions

As a follow-up, we evaluated the community’s perception of our findings. The goal of the follow-up study was to understand to what extent the interpersonal challenges investigated in this study persist in impacting the feeling of welcomeness in OSS in 2024. We conducted an online survey with OSS contributors, which we advertised on social media and the authors’ professional network. We conducted the survey in July 2024. We received 28 responses from OSS contributors from all regions (North America, Central/South America, Europe, Asia Pacific, and Africa). We also conducted an audience poll during the keynote at the 2024 SciPy Conference [69], which is dedicated to scientific computing through open-source Python software. We received an additional 25 responses.

In the surveys, we presented the seven interpersonal challenges and asked if they still happen today and to what extent they can influence how a contributor feels welcome in the OSS community. Additionally, we summarized our results into seven statements and asked the respondents’ agreement level. Finally, we added an open question so they could share additional thoughts about challenges and feelings of being included in OSS. All questions were optional. The questionnaire and its answers are part of the replication package [46].

## IV. ANALYSES AND RESULTS

In this section, we assess each research question, followed by a composite assessment of the answers to the open question about the reasons to (not) feel welcome.

#### A. How are interpersonal challenges associated with the feeling of welcomeness in OSS?

In this section, we describe our results for RQ1, which include the validate the measurement model (Sec. IV-A1), and then we evaluate the hypothesis (**H1**. Experiencing interpersonal challenges is negatively associated with feeling welcome) and control variables in the structural model (Sec. IV-A2), both computed through the survey data.

We assess the significance of our model using the Structural Model Evaluation - Partial Least Squares (PLS-SEM) and following the evaluation protocol proposed by previous research [63, 70] to make results consistent with our claims. The path weighting scheme was estimated using SmartPLS 4 [71].

PLS-SEM is widely used in software engineering research to explore relationships between theoretical concepts not directly observable but measured through indicators or manifest variables [70]. In our study, this applies to interpersonal challenges. As in related work II, PLS-SEM constructs a “proxy” of manifest variables as weighted composites and tests their association with another variable—welcomeness.

1) *Measurement Model Evaluation*: We conducted a series of tests to validate the measurement of the theoretical concept, including convergent validity, internal consistency reliability, discriminant validity, and collinearity.

Convergent validity relates to the degree to which a measure correlates positively with alternative measures of the same construct. Our model contains one reflective latent variable (Interpersonal Challenges). Changes in the latent construct should be reflected in changes in the indicator variables [63]. We used two metrics to assess convergent validity: the Average Variance Extracted (AVE) and the loading of an indicator onto its construct (the outer loading) [63]. The AVE value for the latent construct in our model is 0.52, which is above the threshold of explainability (0.5). For the outer loading, we remove variables that does not reflect a sufficiently large impact in the latent variable. An outer loading of 0.708 is considered sufficient (because  $0.708^2 \approx 0.50$ , which means at least half of the indicator’s variance can be explained by the latent variable), and 0.60 is considered sufficient for exploratory studies [63]. All indicators of our latent construct exceeded this threshold, ranging between 0.62 and 0.83.

We also verified the Internal Consistency Reliability, or how well the different indicators for interpersonal challenges are consistent with one another and able to reliably and consistently measure the constructs. We calculated both Cronbach’s  $\alpha$  and Composite Reliability (CR) to measure the reliability [63]. A desirable range of values for both Cronbach’s  $\alpha$  and CR is between 0.7 and 0.9 [63]. The Cronbach’s  $\alpha$  was 0.871, and CR was 0.887 for our model.

We also verified whether each construct represented different concepts or entities through discriminant validity tests to assess the discriminant validity utilizing the Heterotrait-monotrait (HTMT) ratio of correlations [72]. The discriminant validity could be problematic if the HTMT ratio exceeds 0.9 [72]. The HTMT ratio between the latent construct and the endogenous variable was 0.460, which is not a significant threat to this study.

Finally, we evaluated if the indicators from the exogenous construct (Interpersonal Challenges) were independent. We calculated their collinearity using the Variance Inflation Factor (VIF). A widely accepted cut-off value for the VIF is 5 [63]; in our model, all VIF values are below 2.

Given these results, we argue that the model is valid for evaluating our hypothesis.

2) *Evaluating the Hypothesis in the Structural Model*: We found support for our hypothesis ( $p = 0.000$ ): experiencing interpersonal challenges is negatively associated with feeling welcome. This relationship is depicted by an arrow in the diagram in Fig. 2. The path coefficient, interpreted as a standardized regression coefficient, indicates the direct effect of one variable on another. With a negative path coefficient ( $B = -0.493$ ), we conclude that interpersonal challenges are negatively associated with the feeling of being welcome. This means that an increase of one standard deviation in Interpersonal Challenges results in a decrease of 0.493 standard deviations in Feeling Welcome.

We assessed the relationship between constructs and the predictive capabilities of the theoretical model. The  $R^2$  value of the endogenous variable in our model (feeling welcome) was 0.301. While some scholars have suggested thresholds to evaluate such values, there is considerable debate about setting such thresholds. Other factors might play a role in feeling welcome. We consider the  $R^2$  value low-moderate.

We analyzed the Standardized Root Mean Square Residual (SRMR), which is a common fit measure to detect misspecification of PLS-SEM models [70]. SRMR is a model fit measure that quantifies the divergence between observed and estimated covariance matrices. The literature suggests using SRMR to compare and evaluate SEM models [73]. SRMR is the square root of the sum of the squared differences between the model-implied and the empirical correlation matrix, or the Euclidean distance between the two matrices [74]. A value of 0 for SRMR would indicate a perfect fit, and values less than 0.08 (conservative) or 0.10 (more lenient) are considered a good fit [73]. Our results suggest a good fit of the empirical data in the theoretical model ( $SRMR = 0.075$ ).

*Control Variables*: We also examined our data to determine if being part of gender or race minorities or having disabilities could strengthen or weaken feeling welcome. We found that identifying as gender or race minorities is associated with not feeling welcome, but we did not find a significant association between having disabilities and feeling welcome.

Table II shows the results of the analysis, including the mean of the bootstrap distribution ( $B$ ), the standard deviation ( $SD$ ), the 95% confidence interval, and p-values.

TABLE II: Standardized path coefficients ( $B$ ), standard deviations (Std), confidence intervals (95% CI), and p-values.

	$B$	Std	95% CI	p-value
H1	<b>-0.49*</b>	0.035	(-0.56, -0.42)	<0.001
Gender Minorities → Feeling Welcome	<b>-0.25*</b>	0.078	(-0.41, -0.10)	0.001
Race Minorities → Feeling Welcome	<b>-0.16*</b>	0.073	(-0.31, -0.02)	0.028
Disabilities → Feeling Welcome	-0.02	0.085	(-0.19, 0.14)	0.779

Coefficients marked with \* are statistically significant

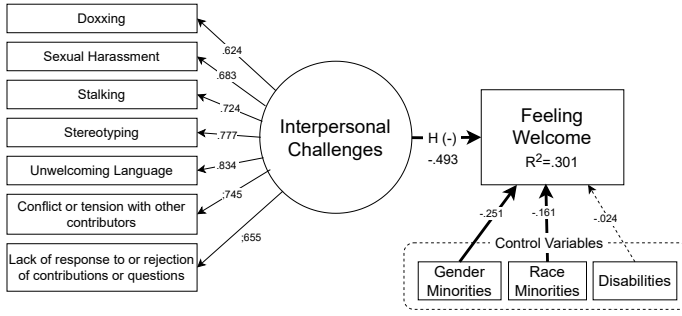


Fig. 2: Item loadings and path coefficient ( $p < 0.05$  indicated by a full line). Non-significant links are indicated with a dashed line.

**RQ1.** Experiencing interpersonal challenges is associated with not feeling welcome within OSS projects. Gender and racial minorities feel less welcome than their counterparts.

*B. How does the association between interpersonal challenges and feeling welcome differ for those who are in gender, race, or disability minorities?*

We answered RQ2 by performing a Multi-Group Analysis in PLS-SEM to investigate the moderation of the theoretical model across demographic groups. We used MGA to examine how the results from RQ1 vary across cohorts.

We started by using the groups of (i) gender, (ii) race, and (iii) disabilities (see Table I). Following, we observed that configural invariance is established in our model as no different settings or treatments were applied to the groups. Still, in this step, we ran the permutation test in SmartPLS and verified that compositional invariance is established for all latent variables in the PLS path model. We established partial measurement invariance, and MGA is suitable for the examined groups [75].

Finally, we analyzed the differences between the coefficients' paths for the groups. If they are significant, they can be interpreted as having moderating effects. In this step, we used parametric tests to assess the multi-group differences, as in Step 2 we established partial measurement invariance [75]. As Table III shows, parametric tests showed statistical differences regarding gender and disabilities but not regarding race. Gender minorities had ( $B=-0.55$ ) a 1.6x stronger association between interpersonal challenges and not feeling welcome than men ( $B=0.35$ ). Parametric tests showed no significant differences for race or disabilities.

TABLE III: Multi-Group Analysis per demographics

Group	Gender		Race		Disabilities		All
	Minorities	Men	Minorities	White	Yes	No	
Sample size (N)	167	539	220	486	547	157	706
Feeling Welcome ( $R^2$ )	0.31	0.12	0.24	0.31	0.27	0.16	0.31
H1	<b>-0.55*</b>	<b>-0.35*</b>	-0.51*	-0.55*	<b>-0.52*</b>	<b>-0.40*</b>	-0.49*

Coefficients marked with \* are statistically significant, and in gray are statistically significant differences between groups

**RQ2** Interpersonal challenges were more strongly associated with the lack of feeling welcome for gender minorities and people with disabilities, while no significant difference was observed for race minorities in this context.

*C. How differently do those in gender, race, or disability minority face interpersonal challenges?*

In this section, we investigate the frequency of interpersonal challenges experienced by the three demographic groups of gender, race, and disability. In Fig. 3, we present the answers to the questions about the frequency of experiencing interpersonal challenges in OSS (from left to right: Frequently, Occasionally, Rarely, and Never). We analyzed the graph by comparing the percentage of answers in the two top frequencies (Frequently and Occasionally). After each item, we show the delta, comparing the percentages for each minority and its counterpart.

Next, we conducted an ordinal logistic regression to follow up on the differences observed in the visualization (Table IV). We have checked our variables' covariance matrix and VIF to confirm no multicollinearity. We used the Benjamini-Hochberg approach to adjust the p-value for the coefficients [76] to control the false rate due to multiple comparisons. As we analyzed interpersonal challenges using three demographic variables as independent predictors, resulting in multiple comparisons, we applied the Benjamini-Hochberg approach to adjust the p-values in our ordinal logistic models to control the false discovery rate (FDR) given the multiple hypothesis tests. This method allows control of FDR while maintaining statistical power and without being overly conservative.

An odds ratio  $> 1$  suggests higher odds for minorities to face the challenge more frequently than their counterparts.

TABLE IV: Odds ratios of challenges per demographics

Challenges	Race Minorities	Gender Minorities	Disability
Lack of Response or Rejection of Contributions	0.87	<b>1.57*</b>	<b>1.68*</b>
Conflict with Other Contributors	0.79	<b>1.58*</b>	<b>1.95*</b>
Unwelcoming Language	0.85	<b>3.09*</b>	<b>2.14*</b>
Stalking	1.21	<b>4.36*</b>	<b>1.97*</b>
Sexual harassment	1.34	<b>6.17*</b>	<b>2.67*</b>
Stereotyping	1.38	<b>3.25*</b>	<b>1.97*</b>
Doxxing	1.67	<b>2.03*</b>	<b>2.67*</b>

Coefficients marked with \* are statistically significant

Gender minorities and people with disabilities from our sample were more likely to experience all interpersonal challenges than their counterparts (see Table IV and Fig. 3). For example, gender minorities were 3.09 times more likely to experience UNWELCOMING LANGUAGE (with a delta of 31.7% for the highest frequency answers) and 6.17 times more likely to experience SEXUAL HARASSMENT (with delta of 22.5%) than men. People with disabilities reported 2.14 more odds of experiencing UNWELCOMING LANGUAGE and 2.67 more odds of experiencing DOXXING than their counterparts.

However, when we look at race minorities, we could not find a statistically significant difference in the odds of experiencing any interpersonal challenge compared to those who identified as white only. From the visualizations (Fig. 3, we can see that the frequency for all the categories does not vary significantly for any of the items and that for the CONFLICT OR TENSION WITH OTHER CONTRIBUTORS challenge, the number of high-frequency answers is slightly higher for the non-white sample.

From a cross-demographics perspective, we observed in Fig. 3 that the number of respondents who reported LACK OF RESPONSE OR REJECTION OF CONTRIBUTIONS *frequently* or *occasionally* is high cross all demographics, noted by



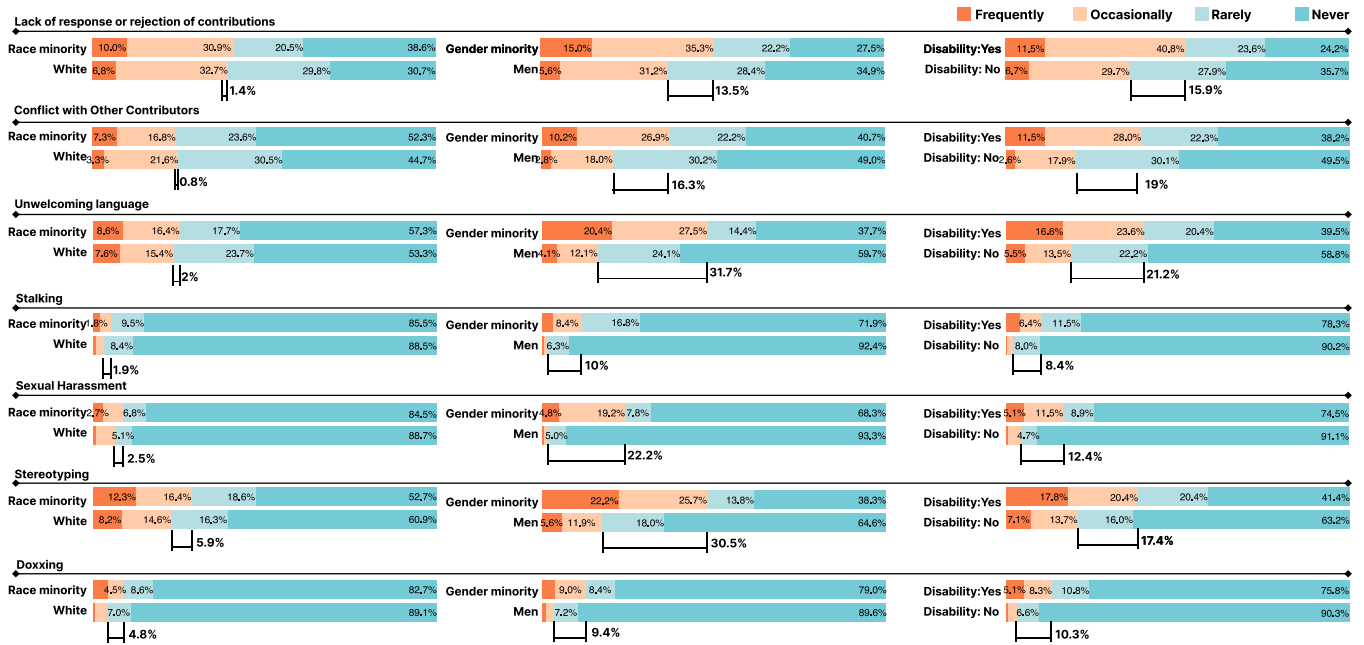


Fig. 3: Responses to the 4-points Likert-scale items for “frequency of interpersonal challenges” (Left:Race; Middle: Gender; Right: Disabilities). The percentage at the bottom of each pair represents the delta, comparing the answers to frequently and occasionally. The whiskers present the differences experiencing challenges (frequent + occasional) between the minority and majority groups.

around half of the respondents. Even with the prevalence of this challenge, gender minorities and people with disabilities reported a higher frequency of occurrence (gender minorities 57% more likely to face the challenge more often than their counterparts, and people with disabilities 68% more likely).

Another major concern for contributors is the CONFLICT OR TENSION WITH OTHER CONTRIBUTORS. While we did not observe any differences regarding race, gender, minorities and people with disabilities reported higher odds (1.58 and 1.95, respectively) of experiencing this challenge than their counterparts. This suggests that these two groups could experience unique challenges related to conflicts. We noticed that challenges related to UNWELCOMING LANGUAGE and STEREOTYPING followed a similar trend, as conflicts. To further investigate this, we conducted a follow-up analysis to assess the correlation of CONFLICT OR TENSION WITH OTHER CONTRIBUTORS with UNWELCOMING LANGUAGE and with STEREOTYPING for the minority groups. We noted that the correlations were statistically significant, with (a least) moderate correlation (Spearman’s rank correlation coefficient  $\rho$  [77]) — 0.71, 0.67, 0.69 for UNWELCOMING LANGUAGE/CONFLICT and 0.59, 0.59, 0.64 for STEREOTYPING/CONFLICT, considering gender, race, and disabilities minorities, respectively. This may suggest that the conflicts may have a relationship with these other two challenges, especially for gender minorities and contributors with disabilities.

Both Fig. 3 and Table IV highlight alarming concerns regarding SEXUAL HARASSMENT, STALKING, and DOXXING, which present the highest odds ratio disadvantaging minorities. Even though these interpersonal challenges are less common than others, they bring a non-negligible delta (when we compare the highest frequency answers from gender and disabilities

minorities with their counterparts), and a *really low* number of “frequently” and “occasionally” answers of the majorities. By looking at Table IV, we observe that gender minorities are 6.17 times more likely to experience sexual harassment than men, with high odds ratios also for STALKING (4.36) and DOXXING (2.03); these interpersonal challenges are at least two times more likely to be experienced by gender minorities than by men.

**RQ3** Gender minorities and people with disabilities face significantly more interpersonal challenges compared to their counterparts. The lack of response or rejection of contribution is prevalent among both minorities and majorities, while sexual harassment, stalking, and doxxing are prevalent among minorities. Conflict or tension with other contributors is experienced together with either unwelcoming language or stereotyping by minorities of gender, race, and disabilities.

#### D. Composite Assessment

As part of our mixed method concurrent embedded strategy [19], we had a composite assessment from the open question of reasons to (not) feel welcome. During this strategy, we qualitatively analyzed data to reside side by side with quantitative analysis performed to answer RQ1, RQ2, and RQ3, having different pictures to provide an overall composite assessment of the problem, relieving details about why our participants feel (or do not feel) welcome in OSS.

People are often biased against others outside of their social group, showing prejudice (emotional bias), STEREOTYPES (cognitive bias), and discrimination (behavioral bias). Both minorities and majorities mentioned being discriminated against due to their demographic characteristics of race, disabilities, gender, age, religion, political ideas, or sexual preferences.



For example, race minorities feel discriminated against during in-person gatherings. P80 mentioned being “*always the only African American and people act like: why are you here?*”. During technical group discussion meetings, P141 remembered that “*being African American and the white people would all sit on the other side of the room.*” Discrimination against gender minorities was mostly mentioned as a glass ceiling, as a woman reported “*they are still excluded from the critical discussions.*” (P508). Discrimination against majorities appeared as another side of the spectrum. In a notable contrast, men, white, middle-aged, or respondents who have a straight sexual orientation feel “*forced into a box* (P123), discriminated against and hated because they are not part of minorities. A white man mentioned “*in the last few years, there has been a dramatic increase in the dismissal of straight, white, male contributions and demonizing of the same group.*” (P431).

Contributors mentioned feeling welcome because they can “*hide behind a profile picture*” and be anonymous (P75, P75, P206, P567). However, when their identities are revealed (DOXXING), P551 complemented: “*you’re dead in the water*”.

Although majorities feel discriminated against, diversity incidents against minorities are still pervasive with “*tons of microaggressions, macroaggressions and SEXUAL HARASSMENT*” (P232). Microaggressions were mentioned as other kinds of CONFLICTS OR TENSION WITH OTHER CONTRIBUTORS who can make “*harsh criticism and impersonal responses*” (P71). During conflicts, P8 misses allies because “*when treated poorly, [she does] not have a single memory of another person in a thread stepping up to disagree*”.

Conflicts were mostly mentioned during communication, and respondents reflected on toxic communication during technical discussions, attributing it to the perception of UNWELCOMING LANGUAGE as being called as *tranny* or a *dumb faggot*” (P8), which leads to unsafety to ask questions (P54), as when asking a question, they faced hostility in which “*people are too negative when they should be inquisitive or just technical*” (P133).

Besides toxicity, contributors feel unwelcome due to LACK OF RESPONSE TO OR REJECTION OF CONTRIBUTIONS OR QUESTIONS when they feel maintainers ignore their attempts of making contact (P105, P168), bringing the perception that “*pull-requests will never get merged*” (P526).

On the positive side, our respondents mentioned feeling welcome due to EXCHANGING SUPPORT when “*people are helpful and share their knowledge*” (P128), but P148 they pondered there is an “*insular culture*” where they need to “*know insiders*” to feel welcome. P282 agreed that “*in the communities that know [him], [he] feels very welcome.. in areas where [he is] unknown, [he] feels like [he does] not have a voice/need to prove [him]self.*”

Although we were interested in interpersonal-related reasons for (not) feeling welcome, another type of reason appeared from the answers. Respondents mentioned feeling unwelcome due to LACK OF ACCESSIBLE TOOLS, particularly for people with sensory impairments like blindness. P518 reported missing accessible Linux mobile phones and remembered an episode when they “*asked a PinePhone developer on Mastodon if*

*accessibility was considered, but [were told that] Linux Desktop accessibility had to develop further, and that [they] should be the change [they] want to see in the world. [They] felt very unwelcome, since [they] know [they] can never come to the knowledge needed to fix an accessibility stack*”.

#### E. Follow-up: Community Perceptions

The follow-up survey was designed to provide insights into how the community perceives our findings (see Sec. III-G). We first presented the seven interpersonal challenges and asked to what extent these challenges influence the feeling of welcomeness. Out of the 53 respondents, 39 answered the question. Among them, 89% reported that the seven interpersonal challenges we investigated continue to significantly or extremely affect welcomeness in OSS. None of the respondents believed these challenges had any effect, indicating that the issues identified in the 2021 survey persisted in 2024 when we conducted the follow-up survey.

Fourteen respondents evaluated a set of statements about how much our findings hold true in 2024. The results, presented in Fig. 4, show that more than 75% of the respondents believe that six out of the seven findings remain completely or largely relevant in 2024 (see numbers on the left of the bars). The finding with the least support refers to the majority groups perceiving D&I initiatives as segregative instead of constructive.

We also invited participants to share their insights about feeling welcome in OSS. One respondent highlighted that stereotyping persists as “*the implicit feeling that people aren’t welcome if they don’t look like 90% of the maintainers.*” Another respondent suggested that future studies should analyze the “*language barrier for non-native speakers*” as an interpersonal challenge affecting the feeling of welcomeness.

In summary, this follow-up study suggests that the findings from the analysis of the 2021 survey remain current at the time this paper was written.

### V. DISCUSSION

We investigated the relationship between *interpersonal challenges* and the *feeling of welcomeness* in OSS projects. Interpersonal challenges are negatively associated with not feeling welcome, even more strongly for gender minorities and people with disabilities. Results align with previous research that discussed that contributors do not feel welcome when facing challenges [6, 13]. In the following, we bring exemplar quotes from the respondents’ responses.

**Interactions and conflicts.** In RQ3, we found a correlation between CONFLICT OR TENSION WITH OTHER CONTRIBUTORS and UNWELCOMING LANGUAGE and between CONFLICT OR TENSION WITH OTHER CONTRIBUTORS and STEREOTYPING, considering the three minority groups of gender, race, and disabilities. Those conflicts can happen through uncivil comments and toxic discussions in mailing lists, comments in pull-requests [29], or also during in-person events.

**The delicate balance of welcoming minorities without appearing unwelcoming to majorities:** Stereotyping is a fixed, oversimplified, and often biased belief about a group of people [56]. Our results from RQ3 showed that gender minorities

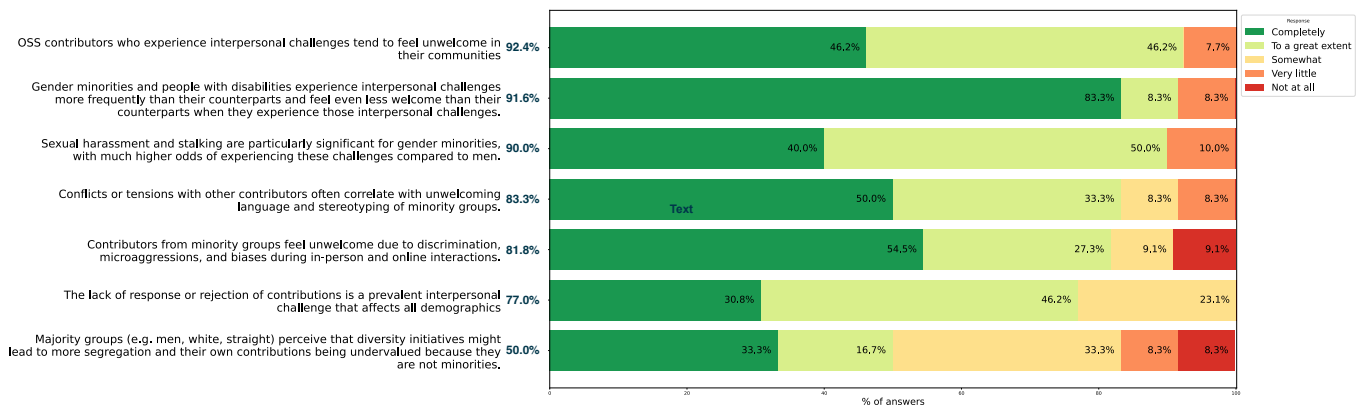


Fig. 4: Results from the follow-up survey. Practitioners answered “How much do you believe these statements hold true in OSS in 2024?” for statements derived from our findings. The list is sorted by the number of respondents who strongly support the statements, i.e., answered “completely” or “to a great extent” (number on the left of the bars). The percentages do not consider the answers marked as “I don’t know.”

and people with disabilities have higher odds of experiencing stereotyping than their counterparts (Table IV). However, in Sec. IV-D, we found that stereotyping hurts regardless if they are part of a minority (e.g., “I feel unwelcome because I’m a woman”) or majority (e.g., “I’m a white guy. Everyone hates me.”). We noticed that minorities and majorities perceive that somehow “diversity is over pushed.” (P56). As noted by P218, “I never felt unwelcome in OSS until this push for inclusion and diversity. I am an Asian female mixed, Greek Yankee and Japanese. I never felt unwelcome in the ‘90s or early 2000s.” Diversity, Equity, Inclusion, and accessibility initiatives must navigate carefully to avoid societal division. While focusing on demographic diversity, they often overlook the deeper cultural shifts required for genuine equity and inclusion. Mere increases in minority representation don’t guarantee progress. True advancement demands transforming organizational culture. To gain public acceptance, we must reject dualistic thinking and emphasize universal inclusion [78].

Our results show that **contributors miss feedback on their questions or acceptance of their contributions**. As Fig. 3 shows, there is a high incidence of people who had experienced LACK OF RESPONSE OR REJECTION OF CONTRIBUTIONS. The reason for this phenomenon can be related to overloaded maintainers with several responsibilities [79, 80], as a result, they have a hard time answering contributors’ requests. More research is required to track the deep roots of why only approximately 1/3 of the respondents reported that they never faced this challenge. Although communities can monitor pull requests and mailing lists to guarantee that members’ posts are not being missed [29], more maintainers are needed to share their workload. Generative AI can be instrumental in providing better feedback and answers to contributors [81, 82].

**The trifecta impeding inclusion: contribution-bias, identity hiding, and doxxing:** Diversity and inclusion initiatives are increasingly implemented in OSS projects [25], but people’s culture still needs to change so minorities are truly included, as mentioned by one of the respondents: “anti-women attitudes are rampant in popular open source projects, it is completely unacceptable, and yet it is tolerated”. When knowing that

discrimination is still tolerated, gender minorities often opt to hide their identities and create pseudonyms to avoid judgment [17] to avoid discrimination and code review bias [32, 83–85]. This behavior was also observed by Ford et al. [86] in online communities, where participants use a “gender neutral alias for websites like technical communities, because [they] get better help when asking questions or answering them.” P1087 (men) claimed for “more anonymity, not judging people based on their irrelevant features.” However, one’s identity should not be considered irrelevant. As we showed in Fig. 3, DOXXING was more frequently experienced by gender minorities than men in our sample. Although people may perceive one’s identity as not relevant, when not revealing their identity, gender minorities may suffer an even more severe disadvantage in survival probability [85]. Although it prevents discrimination, hiding identity can lead to a lack of trust and ultimately cause a higher exit rate for such users [1].

**The pervasive set of challenges still affecting gender minorities** SEXUAL HARASSMENT, [52, 87], STALKING, and STEREOTYPING were the top three interpersonal challenges with higher odds to be experienced by gender minorities (see Table IV. Although most communities have strict rules against negative behaviors, the enforcement of the code of conduct is crucial, as is the need for proactive measures to detect and prevent such challenges from adversely affecting contributors. The code of conduct comprises the collective norms of a community, as mantras that shape the culture of collaboration as well as the community’s expectations and values to create a friendly and welcoming community [55]. Communities can use mining tools to identify gender pronouns in messages such as mailing lists, pull requests, and code reviews. Actions may include providing online training with practical examples of acceptable and non-acceptable behaviors and training for allies to speak up and act as “collaborators, accomplices, and co-conspirators” [88]. Additionally, lawyers should be involved in creating an enforcement plan and adding the appropriate terms to provide transparency about the punishments for those who violate the code of conduct.

**Implications to practice:** The insights from our research

offer practical guidance for OSS communities striving to become more inclusive. For instance, our findings on the particular challenges faced by gender minorities can inform the development of targeted mentorship programs, codes of conduct, and communication guidelines that specifically address these issues. Similarly, the unique experiences of racial minorities and people with disabilities in OSS highlighted in our study could lead to the implementation of community guidelines that promote respect and understanding across diverse backgrounds. By focusing on reducing interpersonal challenges, it is possible that the contributors feel more welcome, and, ultimately, OSS projects can enhance their attractiveness to a broader range of contributors, enriching the community and the software products they develop.

Our findings can inform communities in prioritizing interventions such as training, mentorship, and support programs. For example, since interpersonal challenges were more strongly associated with the lack of feeling welcome for gender minorities and people with disabilities (RQ2), communities aiming to improve the feelings of welcomeness of gender minorities should combat interpersonal challenges, especially sexual harassment, stalking, and doxxing, which our data (RQ3) showed are currently prevalent among minorities. Potential actions can be to have an enforceable Code of Conduct, promote allies who can voice concerns without direct retribution, and set up Diversity, Equity, and Inclusion committees.

**Implications to research:** While our study sheds light on critical aspects of diversity in OSS, it also opens avenues for further research. Future research could explore longitudinal trends in the experiences of minority groups within OSS communities, providing insights into how inclusivity efforts evolve and their long-term effectiveness. Examining the intersectionality of various demographic factors could offer a more comprehensive understanding of the challenges faced by individuals belonging to multiple underrepresented groups. Investigating how different strategies already in place (e.g., code of conduct, near-peer mentoring, demographic-specific support groups) impact the perception of welcomeness by conducting case studies and in-depth qualitative investigations. Another interesting future work would assess the association of welcomeness and sense of belonging—given their theoretical relationship [35].

## VI. THREATS TO VALIDITY

We analyzed the answers to an existing survey conducted by the Linux Foundation Research. The survey was more extensive than what was analyzed in the context of this paper. Another option would be conducting another survey. We opted to use the existing dataset given the breadth and the number of answers of this one, which would not be possible since recruitment in software engineering has been an issue recently [89, 90]. The constructs under analysis are captured by the survey, and one of the authors took part in the design of the instrument, which was tailored to existing measurement instruments for the constructs based on prior literature. Our analysis confirmed that the adopted constructs were internally consistent and scored satisfactorily on convergent and discriminant validity tests.

The hypothesis tested in this study proposes an association between different constructs rather than causal relationships, as the present study is a cross-sectional sample study [91]. We acknowledge the limitation that our respondents comprise contributors who are more likely to feel welcome as they dedicated their time to answering the questionnaire, suggesting a response bias. While it is clear that contributors motivated by some intrinsic-social reasons tend to feel more welcome, a theoretical model such as ours cannot capture a complete and exhaustive list of factors. Moreover, feeling welcome was measured on a single question, and the nuanced complexity of this feeling could benefit by having a latent construct variable with multiple questions. Other factors can play a role, and our results represent a starting point for future studies.

The respondent demographics align with the overall distribution of previous measurements of OSS contributors, making this a suitable starting point for understanding the link between interpersonal challenges and feeling welcome. Further replication studies in individual communities can replicate, validate, and extend our theoretical model. According to previous research, most OSS contributors are men, white, and without disabilities, consistent with our sample. The responses were sufficiently consistent to find full support for the hypothesis.

The survey was conducted in 2021 and provides a rich and comprehensive perspective rarely seen in academic studies. However, since 2021, there may have been changes in the challenges OSS contributors face and their perceptions of inclusivity. To address this threat, we conducted a follow-up survey with practitioners (see Sec. IV-E). Eighty-nine percent of respondents confirmed that the challenges we investigated still substantially influence the sense of welcome in OSS as of 2024, when this paper was written. None said these challenges do not influence welcomeness. Future longitudinal research is needed to investigate how our results may change over time as OSS culture slowly changes.

## VII. CONCLUSION

We assessed the interplay between interpersonal challenges and the perception of welcomeness in OSS communities. Our findings indicate a significant negative correlation between interpersonal challenges and feelings of welcomeness, with pronounced effects on gender minorities and people with disabilities. While these results might not surprise practitioners, our study reveals the strengths of the associations between each challenge and different demographic groups. These findings hold significant implications for communities aiming to enhance welcomeness among their members by focusing on the challenges faced by diverse groups. This involves more than just implementing inclusive policies; it requires cultivating a culture that actively opposes bias, champions diversity, and fosters a sense of belonging and respect.

## VIII. ACKNOWLEDGEMENTS

We thank the respondents and The Linux Foundation Research for making the dataset available. This work was supported by the NSF grants 1900903, 1901031, 2235601, 2247929, 2236198, 2303612, 2303042, 2303043.

## REFERENCES

- [1] B. Trinkenreich, I. Wiese, A. Sarma, M. Gerosa, and I. Steinmacher, "Women's participation in open source software: A survey of the literature," *ACM Trans Soft Eng Methodol*, vol. 31, no. 4, 2022.
- [2] R. Nadri, G. Rodríguez-Pérez, and M. Nagappan, "On the relationship between the developer's perceptible race and ethnicity and the evaluation of contributions in oss," *IEEE Transactions on Software Engineering*, vol. 48, no. 8, pp. 2955–2968, 2021.
- [3] H. Carter and J. Groopman, "The linux foundation report on diversity, equity, and inclusion in open source," 2021, [accessed 2023-06-27].
- [4] B. Vasilescu, D. Posnett, B. Ray, M. G. van den Brand, A. Serebrenik, P. Devanbu, and V. Filkov, "Gender and tenure diversity in github teams," in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 2015, pp. 3789–3798.
- [5] A. Bosu and K. Z. Sultana, "Diversity and inclusion in open source software (oss) projects: Where do we stand?" in *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*. IEEE, 2019, pp. 1–11.
- [6] I. Steinmacher, T. Conte, M. A. Gerosa, and D. Redmiles, "Social barriers faced by newcomers placing their first contribution in open source software projects," in *18th ACM conference on Computer supported cooperative work & social computing*, 2015, pp. 1379–1392.
- [7] I. Steinmacher, T. Conte, and M. A. Gerosa, "Understanding and supporting the choice of an appropriate task to start with in open source software communities," in *HICSS'15*. IEEE, 2015, pp. 5299–5308.
- [8] C. Jensen, S. King, and V. Kuechler, "Joining free/open source software communities: An analysis of newbies' first interactions on project mailing lists," in *HICSS 2011*. IEEE, 2011, pp. 1–10.
- [9] C. Hannebauer and V. Gruhn, "On the relationship between newcomer motivations and contribution barriers in open source projects," in *International Symposium on Open Collaboration*, 2017, pp. 1–10.
- [10] G. Pinto, I. Steinmacher, and M. A. Gerosa, "More common than you think: An in-depth study of casual contributors," in *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, vol. 1. IEEE, 2016, pp. 112–123.
- [11] S. Balali, I. Steinmacher, U. Annamalai, A. Sarma, and M. A. Gerosa, "Newcomers' barriers... is that all? an analysis of mentors' and newcomers' barriers in OSS projects," *Computer Supported Cooperative Work*, vol. 27, no. 3-6, pp. 679–714, 2018.
- [12] B. Trinkenreich, K.-J. Stol, A. Sarma, D. German, M. Gerosa, and I. Steinmacher, "Do i belong? modeling sense of virtual community among linux kernel contributors," in *ICSE 2023*. IEEE, 2023.
- [13] M.-A. Storey, A. Zagalsky, F. Figueira Filho, L. Singer, and D. M. German, "How social and communication channels shape and challenge a participatory culture in software development," *IEEE Transactions on Software Engineering*, vol. 43, no. 2, pp. 185–204, 2016.
- [14] G. Rodríguez-Pérez, R. Nadri, and M. Nagappan, "Perceived diversity in software engineering: a systematic literature review," *Empirical Software Engineering*, vol. 26, pp. 1–38, 2021.
- [15] R. Nadri, G. Rodriguez-Perez, and M. Nagappan, "Insights into non-merged pull requests in github: Is there evidence of bias based on perceptible race?" *IEEE Softw.*, vol. 38, no. 2, pp. 51–57, 2021.
- [16] A. Forte and C. Lampe, "Defining, understanding, and supporting open collaboration: Lessons from the literature," *American Behavioral Scientist*, vol. 57, no. 5, pp. 535–547, 2013.
- [17] A. Lee and J. C. Carver, "Floss participants' perceptions about gender and inclusiveness: a survey," in *ICSE 2019*. IEEE, 2019, pp. 677–687.
- [18] The Linux Foundation Research, "Diversity, equity, and inclusion in open source," <https://www.linuxfoundation.org/research/the-2021-linux-foundation-report-on-diversity-equity-and-inclusion-in-open-source?hsLang=en>, 2021, [Online; accessed 2023-07-19].
- [19] J. W. Creswell and J. D. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications, 2017.
- [20] Z. Feng, "The state of survival in oss: The impact of diversity," in *FSE 2023*, 2023, pp. 2213–2215.
- [21] Z. Feng, M. Guizani, M. A. Gerosa, and A. Sarma, "The state of diversity and inclusion in Apache: A pulse check," in *CHASE 2023*. IEEE, 2023, pp. 150–160.
- [22] B. Trinkenreich, M. Guizani, I. Wiese, A. Sarma, and I. Steinmacher, "Hidden figures: Roles and pathways of successful oss contributors," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. CSCW2, pp. 1–22, 2020.
- [23] B. Trinkenreich, M. Guizani, I. Wiese, T. Conte, M. Gerosa, A. Sarma, and I. Steinmacher, "Pots of gold at the end of the rainbow: What is success for open source contributors?" *IEEE Transactions on Software Engineering*, vol. 48, no. 10, pp. 3940–3953, 2021.
- [24] C. J. Mendez, A. Sarma, and M. Burnett, "Gender in open source software: What the tools tell," in *2018 IEEE/ACM 1st International Workshop on Gender Equality in Software Engineering, GE@ICSE, Gothenburg, Sweden, May 28, 2018*, 2018, pp. 21–24. [Online]. Available: <http://ieeexplore.ieee.org/document/8452746>
- [25] M. Guizani, B. Trinkenreich, A. A. Castro-Guzman, I. Steinmacher, M. Gerosa, and A. Sarma, "Perceptions of the state of d&i and d&i initiative in the asf," in *ICSE-SEIS 2022*, 2022, pp. 130–142.
- [26] Y. Huang, D. Ford, and T. Zimmermann, "Leaving my fingerprints: Motivations and challenges of contributing to oss for social good," in *ICSE 2021*. IEEE, 2021, pp. 1020–1032.
- [27] Apache Software Foundation, "2023 diversity, equity, and inclusion report," 2023, accessed: 2024-03-14. [Online]. Available: <https://www.apache.org/foundation/docs/2023DEIRReport.pdf>
- [28] K. Finley, "Diversity in open source is even worse than in tech overall," <https://www.wired.com/2017/06/diversity-open-source-even-worse-tech-overall/>, 2023, accessed: 2024-03-19.
- [29] C. Miller, S. Cohen, D. Klug, B. Vasilescu, and C. Kaustner, "" did you miss my comment or what?" understanding toxicity in open source discussions," in *ICSE 2022*, 2022, pp. 710–722.
- [30] I. Ferreira, J. Cheng, and B. Adams, "The" shut the f\*\* k up" phenomenon: Characterizing incivility in open source code review discussions," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW2, pp. 1–35, 2021.
- [31] W. E. Powell, D. S. Hunsinger, and B. D. Medlin, "Gender differences within the open source community: An exploratory study," *Journal of Information Technology*, vol. 21, no. 4, pp. 29–37, 2010.
- [32] B. Vasilescu, V. Filkov, and A. Serebrenik, "Perceptions of diversity on git hub: A user survey," in *CHASE 2015*. IEEE, 2015, pp. 50–56.
- [33] V. Singh and W. Brandon, "Discrimination, misogyny and harassment: Examples from oss: content analysis of women-focused online discussion forums," in *Proceedings of the Third Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering*, 2022, pp. 71–79.
- [34] V. Singh, B. Bongiovanni, and W. Brandon, "Codes of conduct in open source software—for warm and fuzzy feelings or equality in community?" *Software Quality Journal*, pp. 1–40, 2021.
- [35] C. A. Hailey, "Racialized perceptions of anticipated school belonging," *Educational Policy*, vol. 36, no. 4, pp. 879–910, 2022.
- [36] N. Gill, "The suppression of welcome," *Fennia-International Journal of Geography*, vol. 196, no. 1, pp. 88–98, 2018.
- [37] C. E. Foor, S. E. Walden, and D. A. Trytten, ""i wish that i belonged more in this whole engineering group:" achieving individual diversity," *Journal of Engineering Education*, vol. 96, no. 2, pp. 103–115, 2007.
- [38] T. Bopp, R. Turick, J. D. Vadeboncoeur, and T. J. Aicher, "Are you welcomed? a racial and ethnic comparison of perceived welcomeness in sport participation," *International Journal of Exercise Science*, vol. 10, no. 6, pp. 833–844, 2017.
- [39] N. Pitterson, R. Streveler, and T. C. Douglas, "Helping newcomers feel "at home": Reflections on the start of careers in engineering education research and implications for broadening participation," in *2022 IEEE Frontiers in Education Conference (FIE)*. IEEE, 2022, pp. 1–6.
- [40] L. M. Williams, "Beyond enforcement: Welcomeness, local law enforcement, and immigrants," *Public Administration Review*, vol. 75, no. 3, pp. 433–442, 2015.
- [41] C. K. Wen, P. L. Hudak, and S. W. Hwang, "Homeless people's perceptions of welcomeness and unwelcomeness in healthcare encounters," *Journal of General Internal Medicine*, vol. 22, pp. 1011–1017, 2007.
- [42] L. Gren and P. Ralph, "What makes effective leadership in agile software development teams?" in *ICSE 2022*, 2022, pp. 2402–2414.
- [43] B. Trinkenreich, M. A. Gerosa, and I. Steinmacher, "Unraveling the drivers of sense of belonging in software delivery teams: Insights from a large-scale survey," in *ICSE 2024*. IEEE, 2024, pp. 974–974.
- [44] The Linux Foundation Research, "Survey data — 2021 diversity, equity, and inclusion in open source," <https://data.world/thelinuxfoundation/2021-diversity-equity-and-inclusion-in-open-source>, 2021, [Online; accessed 2023-07-19].
- [45] J. Hair, C. L. Hollingsworth, A. B. Randolph, and A. Y. L. Chong, "An updated and expanded assessment of pls-sem in information systems research," *Industrial management & data systems*, 2017.
- [46] Blind, "Replication package," <https://figshare.com/s/9b0763fac157e95493be>, 2023.
- [47] C. Bembich and E. Bortolotti, "Early childhood services: Improving

- the relational and emotional skills of educators to ensure inclusiveness,” *Early Childhood Services: improving the relational and emotional skills of educators to ensure inclusiveness*, pp. 30–44, 2017.
- [48] B. Stewart, B. Ju, and K. D. Kendrick, “Racial climate and inclusiveness in academic libraries: perceptions of welcomeness among black college students,” *The Library Quarterly*, vol. 89, no. 1, pp. 16–33, 2019.
- [49] T. L. Foundation, “The linux foundation code of conduct,” 2021, [online; accessed 2024-07-02].
- [50] C. Ringle, D. Da Silva, and D. Bido, “Structural equation modeling with the smartpls,” Bido, D., da Silva, D., & Ringle, C.(2014). *Structural Equation Modeling with the Smartpls. Brazilian Journal Of Marketing*, vol. 13, no. 2, 2015.
- [51] S. Dubberley, A. Koenig, and D. Murray, *Digital witness: using open source information for human rights investigation, documentation, and accountability*. Oxford University Press, USA, 2020.
- [52] D. Nafus, “‘patches don’t have gender’: What is not open in open source software,” *New Media & Society*, vol. 14, no. 4, pp. 669–683, 2012.
- [53] M. L. Pittaro, “Cyber stalking: An analysis of online harassment and intimidation,” *International journal of cyber criminology*, vol. 1, no. 2, pp. 180–197, 2007.
- [54] P. Q. Brady and M. R. Nobles, “The dark figure of stalking: Examining law enforcement response,” *Journal of interpersonal violence*, vol. 32, no. 20, pp. 3149–3173, 2017.
- [55] P. Tourani, B. Adams, and A. Serebrenik, “Code of conduct in open source projects,” in *2017 IEEE 24th international conference on software analysis, evolution and reengineering (SANER)*. IEEE, 2017, pp. 24–33.
- [56] C. McCauley, C. L. Stitt, and M. Segal, “Stereotyping: From prejudice to prediction,” *Psychological Bulletin*, vol. 87, no. 1, p. 195, 1980.
- [57] R. Paul, A. Bosu, and K. Z. Sultana, “Expressions of sentiments during code reviews: Male vs. female,” in *2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER)*. IEEE, 2019, pp. 26–37.
- [58] H. S. Qiu, Y. L. Li, S. Padala, A. Sarma, and B. Vasilescu, “The signals that potential contributors look for when choosing open-source projects,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, pp. 1–29, 2019.
- [59] G. Robles, L. Arjona, A. Serebrenik, B. Vasilescu, and J. González-Barahona, “Floss 2013: A survey dataset about free software contributors: challenges for curating, sharing, and combining,” in *Conference on Mining Software Repositories*, 2014, pp. 396–399.
- [60] M. Guizani, A. Chatterjee, B. Trinkenreich, M. May, G. Noa, L. Russell, G. Zambrano, D. Izquierdo-Cortazar, I. Steimacher, M. Gerosa, and A. Sarma, “The long road ahead: Ongoing challenges in contributing to large oss organizations and what to do,” *Computer Supported Cooperative Work*, 2021.
- [61] R. Kaur and K. K. Chahal, “Exploring factors affecting developer abandonment of open source software projects,” *Journal of Software: Evolution and Process*, vol. 34, no. 9, p. e2484, 2022.
- [62] I. Steinmacher, I. Wiese, A. P. Chaves, and M. A. Gerosa, “Why do newcomers abandon open source software projects?” in *Cooperative and Human Aspects of Software Engineering (CHASE), 2013 6th International Workshop on*. IEEE, 2013, pp. 25–32.
- [63] J. F. Hair, J. J. Risher, M. Sarstedt, and C. M. Ringle, “When to use and how to report the results of pls-sem,” *European business review*, 2019.
- [64] K. Albusays, P. Bjorn, L. Dabbish, D. Ford, E. Murphy-Hill, A. Serebrenik, and M.-A. Storey, “The diversity crisis in software development,” *IEEE Software*, vol. 38, no. 2, pp. 19–25, 2021.
- [65] K. K. Silveira and R. Prikladnicki, “A systematic mapping study of diversity in software engineering: a perspective from the agile methodologies,” in *Cooperative and Human Aspects of Software Engineering*. IEEE, 2019, pp. 7–10.
- [66] J. F. Hair Jr, M. Sarstedt, C. M. Ringle, and S. P. Gudergan, *Advanced issues in partial least squares structural equation modeling*. sage publications, 2017.
- [67] J. Henseler, C. M. Ringle, and M. Sarstedt, “Testing measurement invariance of composites using partial least squares,” *International marketing review*, 2016.
- [68] F. E. Harrell, Jr and F. E. Harrell, “Ordinal logistic regression,” *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*, pp. 311–325, 2015.
- [69] SciPy, “Scipy 2024 conference,” <https://www.scipy2024.scipy.org/>, 2024, accessed: 2024-07-22.
- [70] D. Russo and K.-J. Stol, “Pls-sem for software engineering research: An introduction and survey,” *ACM Computing Surveys (CSUR)*, vol. 54, no. 4, pp. 1–38, 2021.
- [71] M. Sarstedt and J.-H. Cheah, “Partial least squares structural equation modeling using smartPLS: a software review,” 2019.
- [72] J. Henseler, C. Ringle, and M. Sarstedt, “A new criterion for assessing discriminant validity in variance-based structural equation modeling,” *Journal of the academy of marketing science*, vol. 43, no. 1, pp. 115–135, 2015.
- [73] J. Henseler, G. Hubona, and P. Ray, “Using pls path modeling in new technology research: updated guidelines,” *Ind. Manag. Data Syst.*, 2016.
- [74] J. Henseler, T. K. Dijkstra, M. Sarstedt, C. M. Ringle, A. Diamantopoulos, D. W. Straub, D. J. Ketchen Jr, J. F. Hair, G. T. M. Hult, and R. J. Calantone, “Common beliefs and reality about pls: Comments on rönkkö and evermann,” *Organ Res Methods*, vol. 17, no. 2, 2014.
- [75] C. M. Ringle and M. Sarstedt, “Gain more insight from your pls-sem results: The importance-performance map analysis,” *Industrial management & data systems*, 2016.
- [76] D. Thissen, L. Steinberg, and D. Kuang, “Quick and easy implementation of the benjamini-hochberg procedure for controlling the false positive rate in multiple comparisons,” *Journal of educational and behavioral statistics*, vol. 27, no. 1, pp. 77–83, 2002.
- [77] P. Sedgwick, “Spearman’s rank correlation coefficient,” *Bmj*, vol. 349, 2014.
- [78] H. H. Friedman and S. Vlady, “The paradox of dei: How lofty ideals became hated,” *North East Journal of Legal Studies*, vol. 44, no. 1, p. 4, 2024.
- [79] E. Dias, P. Meirelles, F. Castor, I. Steinmacher, I. Wiese, and G. Pinto, “What makes a great maintainer of open source projects?” in *ICSE 2021*. IEEE, 2021, pp. 982–994.
- [80] N. Raman, M. Cao, Y. Tsvetkov, C. Kästner, and B. Vasilescu, “Stress and burnout in open source: Toward finding, understanding, and mitigating unhealthy interactions,” in *42nd International Conference on Software Engineering: New Ideas and Emerging Results*, 2020, pp. 57–60.
- [81] J. a. Correia, M. C. Nicholson, D. Coutinho, C. Barbosa, M. Castelluccio, M. Gerosa, A. Garcia, and I. Steinmacher, “Unveiling the potential of a conversational agent in developer support: Insights from Mozilla’s PDF.js project,” in *Proceedings of the 1st ACM International Conference on AI-Powered Software*, ser. AIware 2024. New York, NY, USA: Association for Computing Machinery, 2024, p. 10–18. [Online]. Available: <https://doi.org/10.1145/3664646.3664758>
- [82] S. Abedu, A. Abdellatif, and E. Shihab, “Llm-based chatbots for mining software repositories: Challenges and opportunities,” in *Proceedings of the 28th International Conference on Evaluation and Assessment in Software Engineering*, 2024, pp. 201–210.
- [83] J. Terrell, A. Kofink, J. Middleton, C. Rainear, E. Murphy-Hill, C. Parnin, and J. Stallings, “Gender differences and bias in open source: Pull request acceptance of women versus men,” *PeerJ Computer Science*, vol. 3, p. e111, 2017. [Online]. Available: <https://doi.org/10.7287/peerj.preprints.1733v2>
- [84] A. Kofink, “Contributions of the under-appreciated: gender bias in an open-source ecology,” in *Companion Proceedings of the 2015 ACM SIGPLAN International Conference on Systems, Programming, Languages and Applications: Software for Humanity*, 2015, pp. 83–84.
- [85] E. D. Canedo, R. Bonifácio, M. V. Okimoto, A. Serebrenik, G. Pinto, and E. Monteiro, “Work practices and perceptions from women core developers in oss communities,” in *International Symposium on Empirical Software Engineering and Measurement*, 2020, pp. 1–11.
- [86] D. Ford, M. Behroozi, A. Serebrenik, and C. Parnin, “Beyond the code itself: how programmers really look at pull requests,” in *ICSE-SEIS 2019*. IEEE, 2019, pp. 51–60.
- [87] V. Singh, “Women-only spaces of open source,” in *2019 IEEE/ACM 2nd International Workshop on Gender Equality in Software Engineering (GE)*. IEEE, 2019, pp. 17–20.
- [88] T. M. Melaku, A. Beeman, D. G. Smith, and W. B. Johnson, “Be a better ally,” *Harvard Business Review*, vol. 98, no. 6, pp. 135–139, 2020.
- [89] A. Alami, M. Zahedi, and N. Ernst, “Are you a real software engineer? best practices in online recruitment for software engineering studies,” in *International Workshop on Methodological Issues with Empirical Studies in Software Engineering*, 2024.
- [90] A. Danilova, A. Naiakshina, S. Horstmann, and M. Smith, “Do you really code? designing and evaluating screening questions for online surveys with programmers,” in *ICSE 2021*. IEEE, 2021, pp. 537–548.
- [91] K.-J. Stol and B. Fitzgerald, “The abc of software engineering research,” *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 27, no. 3, p. 11, 2018.