Reinforcing Diversity Company Policies: Insights from StackOverflow Developers Survey

Karina Kohl Silveira[a], Soraia Musse[b], Isabel Manssour[c], Renata Vieira[d] and Rafael Prikladnicki[e]

School of Technology, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre, Brazil

Keywords: Software Engineering, Software Development, Diversity, Data Visualization, StackOverflow.

Abstract: Diversity is being intensively discussed by different knowledge areas of society and discussions in Software Engineering, are increasing as well. There are unconscious bias and lack of representativeness and when we talk about characteristics as ethnicity and gender, to mention a few. How can tech companies support diversity, minimizing unconscious bias in their teams? Studies say that diversity builds better teams and delivers better results, among other benefits. Cognitive diversity is linked to better outcomes, and is influenced by identity diversity (e.g., gender, race, etc.), mainly when tasks are related to problem-solving and prediction. In this work, we are interested in understanding the pain points in software engineering regarding diversity and provide insights to support the attraction, hiring and retention policies for more diverse software engineering environments. StackOverflow is a popular community question&answer forum, with a high engagement of software developers. Yearly, they apply a survey, present straightforward results, and made the anonymized results available for download. So, it is possible to perform additional analysis beyond the original ones. Using data visualization techniques, we analyzed 2018 data implying insights and recommendations. Results show that diversity in companies is not yet a conscious decision-making factor for developers assessing a new job opportunity, and respondents from underrepresented groups tend to believe more they are not as good as their peers. We also propose a discussion about the unconscious bias, stereotypes, and impostor syndrome and how to provide support on that.

1 INTRODUCTION

More than ever, software development is a collaborative task. Software development teams are built on people and, lately, the area is becoming aware of the problem of underrepresented groups, like gender, racial, cultural, etc. As mentioned in Vasilescu (Vasilescu et al., 2014) previous work, gender representation in Science, Technology, Engineering, and Mathematics (STEM) related subjects raises the significant attention of researchers and academics, as well as of policy-makers, all noting a significant under-representation of women. Several companies are doing the challenge of embracing diversity in the workforce as Google (Google, 2018a), Facebook (Facebook, 2018), Microsoft (Microsoft, 2018), and SAP (SAP, 2018).

Software Engineering demands well developed problem-solving skills from developers. Considering Agile Methodologies, for example, the Agile Manifesto (Mike Beedle, 2001), put people as critical assets to better performance on development and delivery, by prioritizing *"individuals and interactions over process and tools"*. Roger S. Pressman (Pressman, 2010) says that the agile philosophy emphasizes individual competence (team members) combined with group collaboration as a critical success factor for the team. In other areas, the impact of diversity is also an object of study. Menard et al. (Menard et al., 2018) presents how to adequately address cultural diversity in organizations regarding practices of information protection. Förster (Förster, 2018) shows the effect of ethnic heterogeneity on electoral turnout in Europe, offering insights into the role ethnic heterogeneity plays in political participation. Atadero et al. (Atadero et al., 2018) discuss how to create an in-

[a] https://orcid.org/0000-0002-2964-4681
[b] https://orcid.org/0000-0002-3278-217X
[c] https://orcid.org/0000-0001-9446-6757
[d] https://orcid.org/0000-0003-2449-5477
[e] https://orcid.org/0000-0003-3351-4916

Silveira, K., Musse, S., Manssour, I., Vieira, R. and Prikladnicki, R.
Reinforcing Diversity Company Policies: Insights from StackOverflow Developers Survey.
DOI: 10.5220/0007707901190129
In *Proceedings of the 21st International Conference on Enterprise Information Systems (ICEIS 2019)*, pages 119-129
ISBN: 978-989-758-372-8

clusive environment in the first years of engineering courses in the light of efforts to broaden the participation of women and people of color in engineering degree programs and careers. Software engineering environments need to be safe for diversity and developers, managers and all other roles in these environments need to understand what are the factors that impact work environments, and to help companies to refine their diversity and inclusion policies.

This paper aims to show our insights, based on an analysis of StackOverflow Developers' Survey data. To achieve that, we are interested in exploring the following research question:

RQ. What companies working with software engineering should focus to attract, hire and retain talents on the specter of diversity?

To answer this question, we derived three other questions:

RQ1. What is being considered by developers when assessing job opportunities?

RQ2. What are their professional objectives for the future?

RQ3. How confident developers are in their programming skills?

Our analysis shows that diversity in the company is not yet a conscious decision-making factor for a developer assessing a new job opportunity except for non-binary and transgenders. Also, respondents that identified themselves as women, non-binary and transgenders tend to doubt more their programming skills and believe they are not as good as their peers than the respondents identified as men. A discussion about the unconscious bias, stereotypes, and impostor syndrome and how to provide support on that is provided in the Results and Discussions in Section 5.

The rest of the paper proceeds as follows. Section 2 presents some background on how some big players in the software engineering industry are dealing with their diversity and inclusion strategies. Section 3 shows some related work regarding data coming from StackOverflow. Section 4 summarizes the methods we used to analyze the data. Section 5 presents our results, Section 7 the threat to validity, and Section 6 discusses them and their implications. Section 8 concludes the paper.

2 BACKGROUND

Page (Page, 2007) says that we cannot tell whether diversity is good or bad unless we first know what diversity is. They characterized diversity as the differences in how people see, categorize, understand, and go about improving the world, recognizing different dimensions of diversity: cognitive and identity. Cognitive Diversity is the difference between how we interpret, reason and solve problems - how we think. Identity Diversity is determined by affiliation with a social group as gender, culture, ethnicity, religion, sexual orientation, etc. Identity diversity and cognitive diversity often go hand in hand. People belonging to different identity groups, or with different life experiences, also tend to acquire diverse cognitive tools. Education, life experiences, and identity can all contribute to cognitive diversity. How much of these matters depend on the task. For identity diversity to be beneficial in a group, it must link with cognitive diversity, and it happens when tasks are related to problem-solving and prediction, so the identities translate into relevant tools. Also, when the members of the group have little or no preference diversity, and when they get along with one another. In these cases, identity diverse groups do perform better than homogeneous groups (Page, 2017).

2.1 Efforts to have Diverse Workforces

To have diversity in the workforce is a challenge embraced by several companies. Table 1 summarizes the Gender distribution and Table 2 summarizes Race/Ethnicity distribution in some huge technology companies.

Google (Google, 2018a) is one of the technology companies that are engaged in increasing their numbers in diversity and inclusion area. They believe once their mission is to organize the world's information and make it universally accessible and useful, it means for every one everyone and, to do that well, a workforce that is a representative of the users they served is important. They present an accelerated approach to diversity and inclusion and share, annually, their Diversity Annual Report on how they plan to deliver their strategy. There is a disclaimer that current gender reporting they published is not inclusive of their non-binary population and they intend to take into account research such as Transgender-inclusive measures of sex/gender for population surveys. One of the discussed points is regarding unconscious bias, and they work with the idea that understanding bias and its intersection with the workplace and the communities around is crucial to promote change. Science shows that everyone is biased, once the human brain is predisposed to negative stereotypes (Spiers et al., 2017) and they do not expect people to rid themselves of all bias, but we want them to recognize it. Research shows that when we are more aware of unconscious bias, we make more objective deci-

sions. To date, 84% of Google's people managers have taken Unconscious Bias training, and they introduced Unconscious Bias workshops into all new employees orientations. Also, Google provides numerous guides to practices and tools to improve their people processes (Google, 2018b).

Facebook (Facebook, 2018) shares, since 2014, their journey to build a diverse company that reflects the global community they serve. For 2018 report they highlight what they believe is working and where they can do better. Facebook believes that diversity is critical to their success as a company once people from all backgrounds rely on Facebook to connect with others, and they will better serve their needs with a more diverse workforce. To attract the best and the brightest, they believe that effective recruiting is critical for building a diverse company. To do that, they work with organizations that support people of color and women in computer science and engineering, some of which include Anita Borg/Grace Hopper (Institute, 2018), SHPE (SHPE, 2018) and NSBE (NSBE, 2018), as well as many others that support a broad range of groups. They see steady increases in hiring rates for underrepresented people since they started testing this approach in 2015.

Microsoft (Microsoft, 2018) says that diversity and inclusion are much more than gender and race demographics. It is about different cultures, religions, ages, political affiliations, education, and sexual orientations. The commitment to diversity and inclusion means creating an environment where everyone feels included and valued. To do that, Microsoft is committed to build and expand the pipeline for diverse technical candidates. They work with girls in early age to sparking their interest in technology careers. Also, work in partnerships with associations for women in STEM and expanding their military academy program to military bases worldwide and still to seek meaningful ways to encourage and cultivate future workforce. Besides, they are also reviewing their traditional recruiting practices to become more expansive and more inclusive in the processes. They expanded the scope of universities where they recruit, such as Historically Black Colleges and Universities (HBCUs) and also have programs for autism hiring. Managers should take Inclusive Hiring training. They believe that build a diverse culture is a critical element to spark innovation and allow unique perspectives and insights to the surface.

SAP (SAP, 2018) presents four diversity areas they work on: Gender, Cross-Generational, Culture & Identity, and Differently Abled. SAP is dedicated to eliminating bias in the workplace and want to enable individuals to be recognized for what they have to contribute. The idea is to embrace and encourage different perspectives and that they are stronger by the unique combination of culture, race, ethnicity, age, gender, sexual orientation, gender identity or expression, physical or mental ability, and work-life situations. It can be highlighted the work cross-generational where people at different stages of life bring a variety of perspectives and experiences to the company — also, differently abled people and the Autism at Work program that leverages the abilities and perspectives of people with autism to foster innovation and to help customers become intelligent enterprises. The program aims to reduce barriers of entry so qualified individuals can fully develop their potential, and it employs over 140 people and in 12 countries.

Table 1: Gender Distribution in Tech Companies (Technical Roles).

Company	Female	Male
Google	21.4%	78.6%
Facebook	21.6%	78.4%
Microsoft	19.0%	81.0%
SAP	33% 1]	-

SAP published only the number of Women in the entire company with no differentiation of kind of role.

Table 2: Ethnicity Distribution in Tech Companies.

Company	Asian	Black	Latinx	White	Other
Google	41.41%	1.5%	2.8%	50.7%	3.8%
Facebook	50.3%	1.3%	3.1%	42.7%	2.6%
Microsoft	38.2%	2.7%	4.3%	52.3%	2.4%

3 CONTEXT SELECTION

StackOverflow[1] is a popular question and answer sites for developers based on gamification where participants earn reputation points and badges that can be seen as a measure of their expertise by peers and potential recruiters and are known to motivate users to contribute more (Vasilescu, 2014). An extensive list of academic papers using StackOverflow data has been published since it was created in 2008 as Vasilescu et al.(Vasilescu et al., 2013), Bosu et al.(Bosu et al., 2013) and, Berger et al.(Berger et al., 2016) to mention a few. However, we did not find any academic paper using their Annual Developers' Survey data (published since 2011).

Vasilescu et al. (Vasilescu et al., 2013) investigated the interplay between StackOverflow activi-

[1]https://www.stackoverflow.com

ties and the development process, reflected by code changes committed to the social coding repository, GitHub. Their study showed that active GitHub committers ask fewer questions and provide more answers than others. They also observed that active StackOverflow askers distribute their work in a less uniform way than developers that do not ask questions. And, finally, they showed that despite the interruptions incurred, the StackOverflow activity rate correlates with the code changing activity in GitHub.

Once earning a high reputation score requires technical expertise and sustained effort, Bosu et al. (Bosu et al., 2013) analyzed the StackOverflow data from four perspectives to understand the dynamics of reputation building on it. The results provided guidance to new StackOverflow contributors who want to earn high reputation scores quickly indicating the following activities to help to build reputation: answering questions related to tags with lower expertise density, answering questions promptly, being the first one to answer a question, being active during off-peak hours, and contributing to diverse areas.

Berger et al. (Berger et al., 2016) studied Question and Answer sites, like StackOverflow, that uses reward systems to incentive users to answer fast and accurately. They investigated and predicted the response time for questions on StackOverflow, that benefit from an additional incentive through so-called bounties. In their findings, they noted that topic related factors provide much stronger evidence than previously found elements for these questions.

Krueger et al. (Krüger et al., 2017) work is about researchers performing empirical studies in the industry to gain qualitative insights into a real-world problem. However, common critics are the diversity and selection process of participants. To address these issues, they propose to improve the integration of question-answering systems into an empirical study. So, they described approaches to conduct studies in such systems, to exemplify corresponding challenges, and to discuss their potential. They held their research on StackOverflow.

Papoutsoglou et al. (Papoutsoglou et al., 2017) proposed a framework that aims to collect online job advertisements from a web source which concerns Information Technology job offers and to extract from the raw text the required skills and competencies for specific jobs. The selected professional networking web source was StackOverflow, and multivariate statistical data analysis was used to test the correlations between skills and competencies in the job offers dataset.

Yin et al. (Yin et al., 2018) described a novel method for extracting aligned code/natural language pairs from StackOverflow. The method is based on learning from a small number of annotated examples, using highly informative features that capture structural aspects of the code snippet and the correspondence between it and the original natural language query.

4 METHODOLOGY

To answer our research questions, we combined data visualization and data analysis techniques. In this section, we present details about the dataset, the data visualization and preliminary quantitative analysis. A more detailed discussion is done in Section 5.

4.1 Data Description

The data provided by StackOverflow (StackOverflow, 2018a) is based on a survey of 101,592 software developers from 183 countries around the world. Accordingly the criteria used by StackOverflow, this number of responses are what we consider "qualified" for analytical purposes based on completion and time spent on the survey; another approximately 20,000 responses were started but not included in the analysis because respondents did not answer enough questions. From the total qualified responses, 67,441 (66.4%) completed the entire survey. The survey was fielded from January 8 to January 28 and the median time spent on the survey for qualified responses was 25.8 minutes, and the median time for those who finished the entire survey was 29.4 minutes. Survey responses that spent less than 5 minutes were excluded from the final sample. Respondents were recruited primarily through channels owned by StackOverflow. Since respondents were recruited in this way, highly engaged users on Stack Overflow were more likely to notice the links for the survey and click to begin it. Respondents who finished the survey were awarded a "Census" badge as a motivation to complete the survey. The data is anonymized and available for download in CSV format, and under the Open Database License (ODbL). In the following section, we present the quantitative results that answer our sub-research questions and support the answer to the primary research question of this work. First, we analyze the aspects considered relevant by developers regarding their interests in job opportunities. Then we see respondents future goals. The third aspect we investigate is confidence. We show these issues as seen across gender, race, and ethnicity. Finally we present

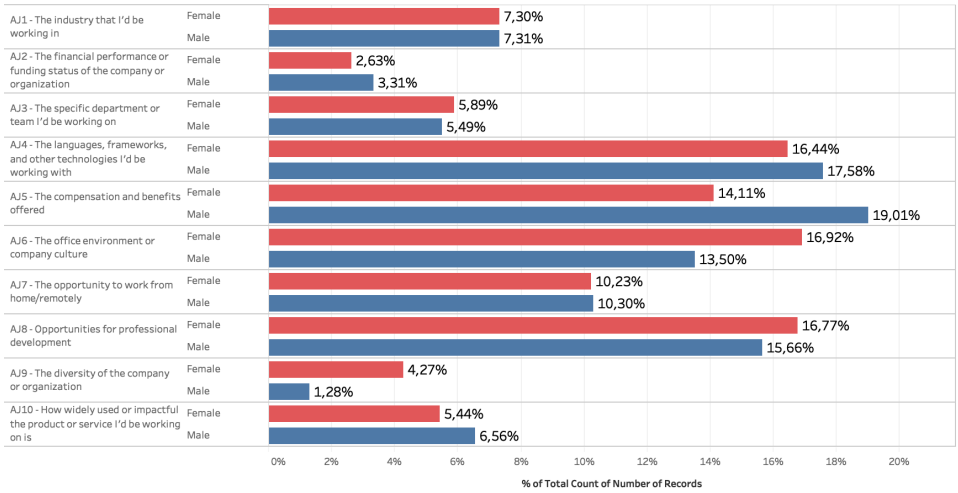

Figure 1: Women/Men Priorities when Assessing a Job.

lessons learned that might serve as input for companies hiring policies.

5 RESULTS

The primary design goal for visualization is to effectively communicate a thorough understanding of the data it represents. This utility of visualization does include usability goals but ultimately revolves around the visualization's ability to help people better understand data (Saket et al., 2016). Nowadays many tools are allowing to visualize and gain faster insights about the data. In this work, we chose to use Tableau Desktop (Tableu, 2018) as the tool to support our visual data analysis.

In the following sections, we present the data visualization that aims to provide support to answer the research questions. In this work, we are not considering salary as a parameter to the discussion because it is a very well discussed subject in the industry and hard news. There are studies about gender pay gaps in different areas and technology as well (Florentine, 2018)(Tarr,)(Ismail, 2018) (Martinson, 2018)(Orphanides, 2018).

5.1 Aspects when Assessing Job Opportunities

This section aims to support to answer our RQ1: *"What is being considered by developers when assessing job opportunities?"*. In the survey, StackOverflow asked the respondents to rank ten aspects when assessing a potential job opportunity. They should rank from 1 (the most important) to 10 (the least important) the following aspects:

- The industry that I'd be working in;
- The financial performance or funding status of the company or organization;
- The specific department or team I'd be working on;
- The languages, frameworks, and other technologies I'd be working with;
- The compensation and benefits offered;
- The office environment or company culture;
- The opportunity to work from home/remotely;
- Opportunities for professional development;
- The diversity of the company or organization;
- How widely used or impactful the product or service I'd be working on is.

For a better analysis of this topic, we split the answers by Gender and Race/Ethnicity. The results follow:

5.1.1 Gender Analysis

For women, the most important aspect when assessing a job is "The office environment or company culture". The last important one is "The financial performance or funding status of the company or organization". For men, the distribution is quite different. Men tend to consider first "The compensation and benefits offered" and the last priority is "The diversity of the company or organization". Figure 1 compares the distribution for women and men.

When comparing the first and last ones, the office environment or company culture, that is the first

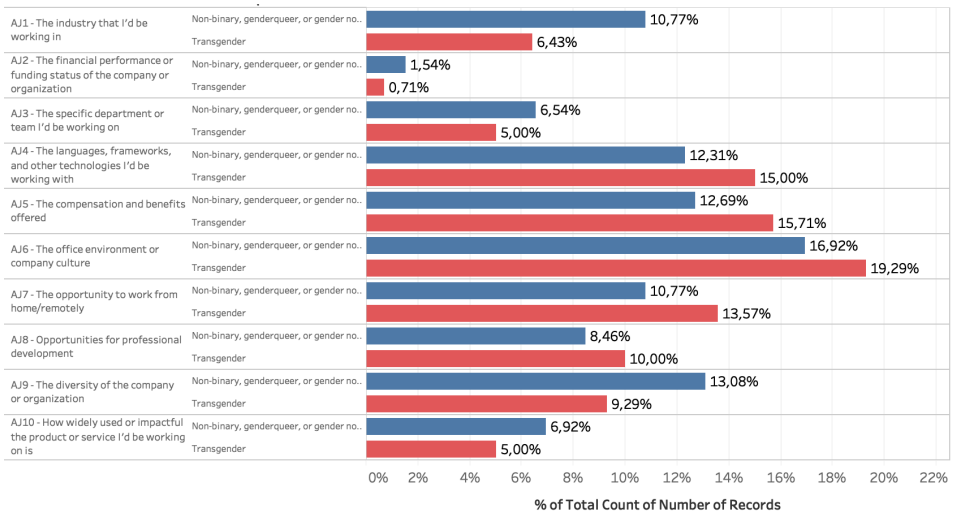

Figure 2: Non-Binary/Transgender Priorities when Assessing a Job.

one for women, appears as fourth for men. The financial performance or funding status of the company or organization appears in the tenth and last place for women and ninth place for men, showing pretty similar importance. The compensations and benefits that appear as first for men are in the fourth place for women. And the diversity of the company, that appears as last important for men appears in ninth place for women. So, even with all the discussions about diversity individual developers that identified themselves as Female or Male are not making it a priority when looking for a job.

However, the panorama changes when we evaluate the data for Non-Binary and Transgender population. Individuals that identified themselves as Non-binaries/Genderqueer or Gender non-conforming put office environment or company culture and diversity of the company as their first and second priorities. Transgenders put culture as the first and diversity as fifth. The financial performance or funding status of the company seems to be very well aligned among all the respondents, no matter the gender - ninth and tenth place as shown in Figure 2.

5.1.2 Race and Ethnicity Analysis

We also analyzed the aspects when assessing a job opportunity by race/ethnicity. The race/ethnicity listed in the StackOverflow survey were: White or of European descent, South Asian, Hispanic or Latino/Latina, East Asian, Middle Eastern, Black or of African descent, Native American/Pacific Islander/Indigenous Australian. Figure 3 compares assessments between Black or Afro Descent and White or European Descent.

Compensation and benefits are the priority for White or European Descent, East Asians, Hispanic/Latins and, Native American/Pacific Islander/Indigenous Australian. Opportunities for professional development are priorities for Black or Afro Descents and Middle Eastern. Languages and frameworks were mentioned as priorities for South Asian. However, diversity of the company is ranked in the tenth for all. Office environment or company culture is indicated in third by East Asians, in fourth by White/European Descents, Native American/Pacific Islander/Indigenous Australian and Middle Eastern and, in the fifth by Hispanic/Latin and Black/Afro Descents.

5.2 Professional Objectives

In this section, the objective is to help to answer our RQ2. *"What are they professional objectives for the future?"*. One of the questions done in the survey was *"What Do Developers Hope To Be Doing in Five Years?"*.

5.2.1 Gender Analysis

While 26% of men want to be Working as a founder or co-founder of their own company, near 16% of women, non-binary and transgenders want the same. Around 42% of women and non-binary wish to be working in a different or more specialized technical role in comparison to approximately 33% of men and transgenders.

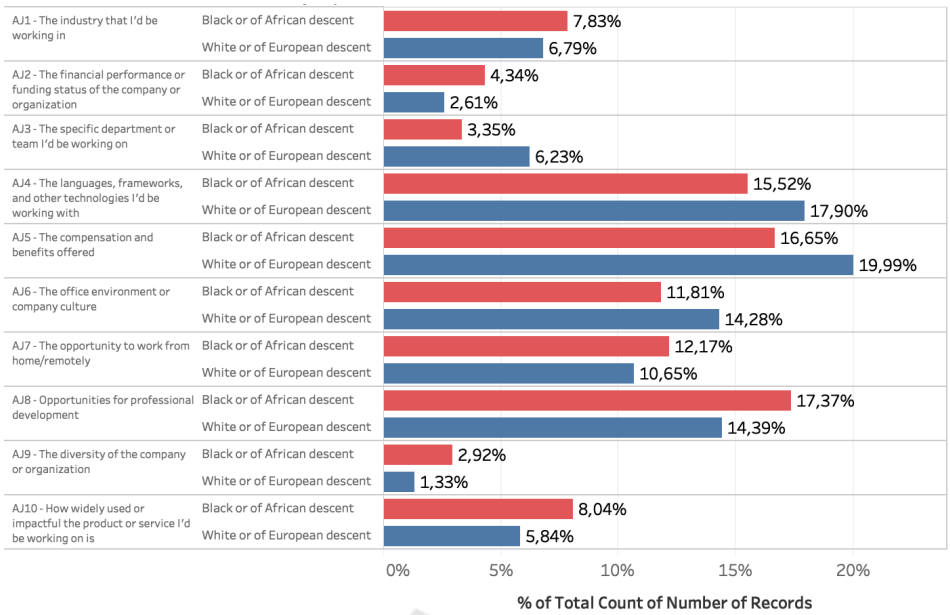

Figure 3: Race/Ethnicity Priorities when Assessing a Job.

5.2.2 Race and Ethnicity Analysis

When analyzing data from the race/ethnicity point of view, we have 44.71% Black or Afro descent wanting to be Working as a founder or co-founder of their own company and 41.53% of South Asians wanting to be Working in a different or more specialized technical role. Figure 5 shows the data for race/ethnicity

5.3 Not as Good Programming as the Peers

This section wants to support to answer our RQ3. *"How confident developers are in their programming skills?"*.

As we can see in Figure 6, most of the answers, no matter the gender, are on the disagreement range or in the neither agree nor disagree. However, men tend to disagree more with the affirmation. On the scale of agreeing and strongly agreeing, we have 29.73% of women that agree or strongly agree that is not as good as their peers, versus 17.28% of men that share the same belief. 24.5% on non-binary and 19.61% of transgenders also believe that are not as good as their peers.

5.4 Insights from Results

In this section we intend to perform a more detailed discussion over the data previously presented and answer our more in-depth research question: *"what companies working with software engineering should focus to attract, hire and retain talents on the specter of diversity?"*

The first analysis presented was about the aspects considered by developers when assessing a new job. Diversity does not seem to be a conscious priority for most of the developers when assessing a new job opportunity. However, non-binaries and transgenders put this on the top of the rank, followed by the culture of the company. Culture appears as first for women and considering gender and race, always mid-ranked. When sharing their results, StackOverflow itself mentioned: *"The tech industry is struggling overall with issues around diversity, and individual developers are not making it a priority when looking for a job."* (Stackoverflow, 2018).

It is essential to recall the case of the Google's employee fired in 2017 after wrote a memo blaming biology for technical's gender gap (Varinsky, 2017). Here we have a mix of two crucial aspects: diversity supported by culture. Even if the diversity of the company is the last item to be considered when a developer considers a new job opportunity, companies may consider this aspect once is linked to enterprise culture and the way to provide a safe work environment for all.

We also evaluated what the respondents intend to be doing in five years. Men want to be owners of their own business when women wish to more specialized technical positions. It sounds that, development programs that offer opportunities for women improve their technical skills may pay off once they

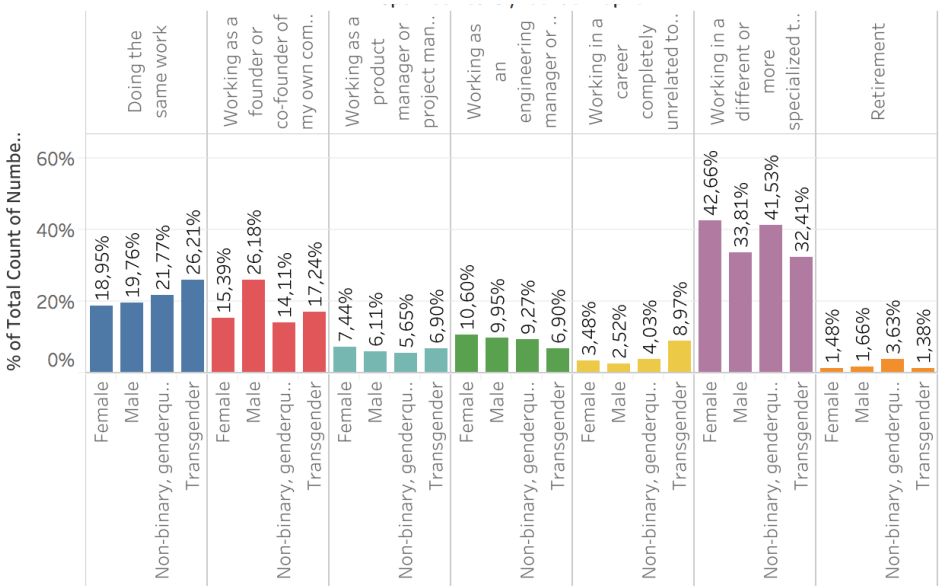

Figure 4: Gender - In Five Years.

demonstrate higher interest than men to be in specialized positions.

The last point analyzed is related to respondents assessing if they believe they are not as good as their peers on their programming skills. Women, Non-binary and transgenders tend to doubt more their programming skills comparing to their peers than men.

6 DISCUSSION

To perform the analysis of these results, it is essential to recall previous events. In 2016, Google published a study about the diversity gap in computer science (Inc. and Inc., 2016). In this study, they identified that male student are more interested and more confident in learning computer science, and that female students rate themselves lower in skills related to Computer Science. Another point identified by the study is that stereotypes may influence implicit beliefs about who can study computer science and might introduce unconscious bias in educators and parents, who may disproportionately and unconsciously encourage students who fit the computer scientist stereotype to pursue Computer Science. For example, male students are more likely than female students to have been told by a teacher (39% vs. 26%) or a parent (46% vs. 27%) that they would be good at Computer Science. Teachers and parents may reinforce stereotypes by telling more male students they think they would be good at Computer Science, thus furthering the underrepresentation of females in Computer Science.

Recent research published in Nature from O'Dea et al. (O'Dea et al., 2018) says that girls are susceptible to conforming to stereotypes in the traditionally male-dominated fields of STEM, and backlash effects hinder girls who try to succeed in these fields. A girl's answer to the question of *"what do you want to be when you grow up?"* will be shaped by her own beliefs about gender, and the collective beliefs of the society she is raised in. The study also compares school grades of girls and boys. They found out that girls tend to earn higher school grades than boys, including in STEM subjects. The prediction of the grade distribution represents that, when all grades are considered, girls on average earn higher grades and are less variable than boys, although there are more highly performing boys than girls at the upper end of the achievement distribution. Therefore, by the time a girl graduates, she is just as likely as a boy to have earned high enough grades to pursue a career in STEM. When she evaluates her options, however, the STEM path is trodden by more male competitors than non-STEM and presents additional internal and external threats due to her and societies' gendered beliefs (stereotype threat and backlash effects). Additionally, the paper says that gender differences in expectations of success can arise due to backlash effects against individuals who defy the stereotype of their gender, or due to gender differences in 'abilities tilt' (having comparatively high ability in one discipline compared to another). Women in male-dominated pursuits, including STEM, face a paradox: if they conform to gender stereotypes, they might be perceived as less

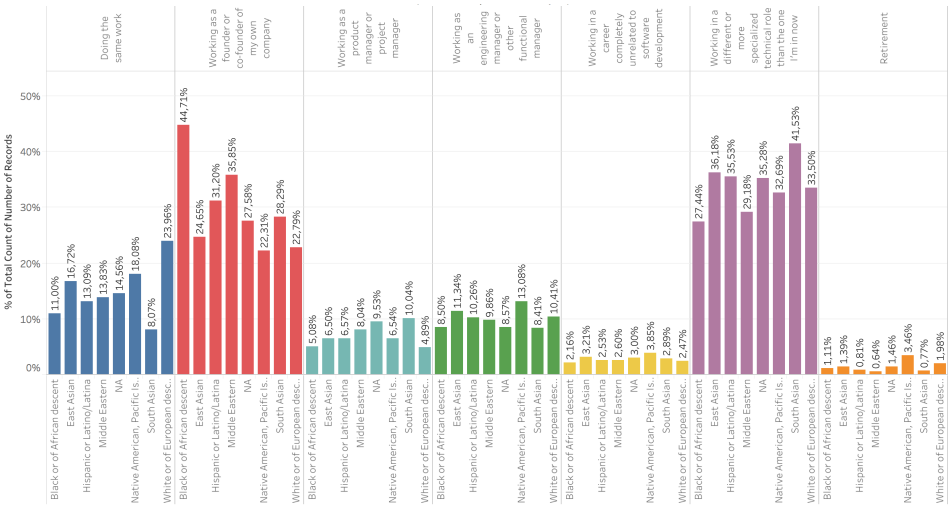

Figure 5: Race/Ethnicity - In Five Years.

competent, but if they defy gender stereotypes and perform 'like a man,' then their progress can be halted by 'backlash' from both men and women.

So, the results observed in the StackOverflow Developers' Survey may be a reflex of stereotypes built from previous personal experiences and be directly related to a confidence gap in women, mainly.

An important point to be considered is what is called the impostor syndrome. Jackson and Heath (Jackson and Heath, 2014) says that impostor syndrome is defined as a psychological phenomenon in which people are unable to internalize their accomplishments. Impostor syndrome affects most people at some point during their careers across all races, all genders, and all ages.

Sukhai (Sukhai and Mohler, 2016) mentions that impostor syndrome is common within the academic environment, particularly at the graduate level student - in STEM fields, mainly, where productivity is a significant measure of a student's success. Impostor syndrome presents itself as a series of feeling or thoughts, and one of them is the frustration with the inability to meet self-set standards (*"I will never be as good as I want to be, so why bother trying?"*). Churchill (Churchill, 2018) points out that over four decades ago when it was given a name by clinical psychologists Pauline Clance and Suzanne Imes in the late 1970s, this feeling was prevalent among high-achieving women.

Follow-on research has shown that impostor syndrome is very real and very prevalent and that its effects are undeniably negative. Therefore also unsurprising that there is a strong correlation between impostor syndrome and anxiety, stress, depression, and burnout, the debilitating condition of exhaustion that can result in talented individuals giving up on promising careers. StackOverflow shares the results about disability status in their results page, with numbers about anxiety, depression and focus (StackOverflow, 2018b), but, unfortunately, the individualized and anonymized results about it were not shared in the CSV file, and we could not make the gender correlation.

To overcome it, it is essential to have a support network that helps to identified impostor syndrome in their workforce. There are some strategies to overcome the syndrome, for example: instruct employees that comparison with others must be done with care (Jackson and Heath, 2014). Comparison without context can be misleading (people will compare themselves with people in another skill level).

That said, what we found in the data from of StackOverflow Developers' Survey suggests the importance of the initiatives to minimize bias and stereotypes that some companies are doing in their hiring process and the process of development of their technical team.

7 THREATS TO VALIDITY

The validity of this work can be subjected to some threats. In the following, threats to internal validity and external validity are illustrated.

External validity refers to how much we can generalize our findings. The presented results are based on data from the StackOverflow community. We suspect that given the high number of respondents and the reputation of the community the results could be generalized outside the scope of our study.

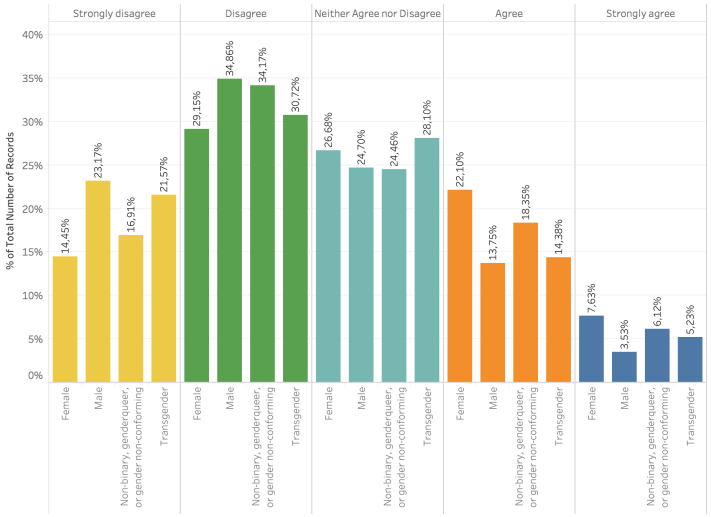

Figure 6: Not as good programming as the peers.

Internal validity often refers to experimenter biases. For our results, the threat of misreading the data visualization analysis and getting to conclusions based on the knowledge areas of the researchers involved in this work.

8 CONCLUSION

The present work set out to provide insights to support the attraction, hiring and retention policies for more diverse and inclusive software engineering environments. Using the anonymized data from StackOverflow Developer's Survey, we performed analysis and correlations beyond their original ones with the support of data visualization techniques that implied in insights to our recommendations. Results show that diversity in the company is not yet a full conscious decision-making factor for developers assessing a new job opportunity, and respondents that identified themselves as women, non-binary and transgenders tend to doubt more their programming skills believing they are not as good as their peers. A discussion about the unconscious bias, stereotypes, and impostor syndrome was done, and we reinforce the importance of initiatives to minimize bias and stereotypes that companies are doing in their hiring process and the process of development of their technical team.

For future work studies, we see opportunities when we selected more specific aspects in the spectrum of the diversity. For example, for cognitive diversity, since there has been an increase in computer science students with the Asperger Syndrom (Ribu, 2010)(Egan, 2005), it is also important to tackle this issue globally in Software Engineering. There is also a need for teaching institutions and software companies to work together to understand these differences better to include them.

ACKNOWLEDGEMENTS

This project is partially funded by FAPERGS, project 17/2551-0001/205-4.

REFERENCES

Atadero, R. A., Paguyo, C. H., Rambo-Hernandez, K. E., and Henderson, H. L. (2018). Building inclusive engineering identities: implications for changing engineering culture. *European Journal of Engineering Education*, 43(3):378–398.

Berger, P., Hennig, P., Bocklisch, T., Herold, T., and Meinel, C. (2016). A journey of bounty hunters: Analyzing the influence of reward systems on stackoverflow question response times. In *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pages 644–649.

Bosu, A., Corley, C. S., Heaton, D., Chatterji, D., Carver, J. C., and Kraft, N. A. (2013). Building reputation in stackoverflow: An empirical investigation. In *Proceedings of the 10th Working Conference on Mining Software Repositories*, MSR '13, pages 89–92, Piscataway, NJ, USA. IEEE Press.

Churchill, E. F. (2018). Is there a fix for impostor syndrome? *Interactions*, 25(3):22–24.

Egan, M. A. L. (2005). Students with asperger's syndrome in the cs classroom. In *ACM SIGCSE Bulletin*, volume 37, pages 27–30. ACM.

Facebook (2018). Facebook diversity update.

Florentine, S. (2018). The hidden truths of it's gender pay gap.

Förster, A. (2018). Ethnic heterogeneity and electoral turnout: Evidence from linking neighbourhood data with individual voter data. *Electoral Studies*, 53:57–65.

Google (2018a). Google diversity.

Google (2018b). Rework guides.

Inc., G. and Inc., G. (2016). Diversity gaps in computer science: Exploring the underrepresentation of girls, blacks and hispanics.

Institute, A. B. (2018). Anita borg/grace hopper.

Ismail, N. (2018). Is there a gender pay gap in the technology sector?

Jackson, D. and Heath, T. (2014). An antidote to impostor syndrome. *XRDS*, 21(2):12–13.

Krüger, J., Schröter, I., Kenner, A., and Leich, T. (2017). Empirical studies in question-answering systems: A discussion. In *Proceedings of the 5th International Workshop on Conducting Empirical Studies in Industry*, CESI '17, pages 23–26, Piscataway, NJ, USA. IEEE Press.

Martinson, J. (2018). Technology will widen pay gap and hit women hardest – davos report.

Menard, P., Warkentin, M., and Lowry, P. B. (2018). The impact of collectivism and psychological ownership on protection motivation: A cross-cultural examination. *Computers and Security*, 75:147–166.

Microsoft (2018). Microsoft global diversity and inclusion.

Mike Beedle, Arie van Bennekum, A. C. W. C. M. F. J. H. A. H. R. J. J. K. B. M. R. C. M. K. S. J. S. D. T. (2001). Agile manifesto.

NSBE (2018). National society of black engineers.

O'Dea, R. E., Lagisz, M., Jennions, M. D., and Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for STEM. *Nature Communications*, 9(1):3777.

Orphanides, K. (2018). Gender pay gap: how big tech companies in the uk compare.

Page, S. E. (2007). *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton University Press.

Page, S. E. (2017). *The Diversity Bonus: How Great Teams Pay Off in the Knowledge Economy (Our Compelling Interests)*. Princeton University Press.

Papoutsoglou, M., Mittas, N., and Angelis, L. (2017). Mining people analytics from stackoverflow job advertisements. In *2017 43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, pages 108–115.

Pressman, R. (2010). *Software Engineering: A Practitioner's Approach*. McGraw-Hill, Inc., New York, NY, USA, 7 edition.

Ribu, K. (2010). Teaching computer science to students with asperger's syndrome. *E. Hjelmås (red.) Norsk Informatikkonferanse*, pages 99–111.

Saket, B., Endert, A., and Stasko, J. (2016). Beyond usability and performance: A review of user experience-focused evaluations in visualization. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*, BELIV '16, pages 133–142, New York, NY, USA. ACM.

SAP (2018). Sap diversity.

SHPE (2018). Society of hispanic professional engineers.

Spiers, H. J., Love, B. C., Le Pelley, M. E., Gibb, C. E., and Murphy, R. A. (2017). Anterior temporal lobe tracks the formation of prejudice. *Journal of Cognitive Neuroscience*, 29(3):530–544. PMID: 27800703.

Stackoverflow (2018). How do developers assess potential jobs?

StackOverflow (2018a). Stackoverflow developers survey 2018.

StackOverflow (2018b). Stackoverflow survey - developer profile disability status.

Sukhai, M. and Mohler, C. (2016). *Creating a Culture of Accessibility in the Sciences*. Academic Press.

Tableu (2018). Tableau.

Tarr, T. By the numbers: What pay inequality looks like for women in tech.

Varinsky, D. (2017). Google - employee fired.

Vasilescu, B. (2014). Software developers are humans, too! In *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW Companion '14, pages 97–100, New York, NY, USA. ACM.

Vasilescu, B., Capiluppi, A., and Serebrenik, A. (2014). Gender, representation and online participation: A quantitative study. *Interacting with Computers*, 26(5):488–511.

Vasilescu, B., Filkov, V., and Serebrenik, A. (2013). Stackoverflow and github: Associations between software development and crowdsourced knowledge. In *2013 International Conference on Social Computing*, pages 188–195.

Yin, P., Deng, B., Chen, E., Vasilescu, B., and Neubig, G. (2018). Learning to mine aligned code and natural language pairs from stack overflow. In *Proceedings of the 15th International Conference on Mining Software Repositories*, MSR '18, pages 476–486, New York, NY, USA. ACM.