2019 IEEE/ACM 12th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)

A Systematic Mapping Study of Diversity in Software Engineering: A Perspective from the Agile Methodologies

Karina Kohl Silveira
School of Technology
PUCRS
Porto Alegre, Brazil
karina.kohl@edu.pucrs.br

Rafael Prikladnicki
School of Technology
PUCRS
Porto Alegre, Brazil
rafaelp@pucrs.br

***Abstract*—Diversity is being discussed intensively by different knowledge areas. Some studies show that diversity builds better teams, delivers better results, and more. Cognitive diversity is linked to better outcomes, and studies show that cognitive diversity is influenced by identity diversity (e.g., gender, race, age, etc.), mainly when tasks are related to problem-solving and prediction. The discussions about diversity in Software Engineering are increasing as well. There is a known lack of representativeness from some groups when we talk about identity diversity as, for example, gender and race. To support diversity in Software Engineering is valuable and it is especially relevant once we are dealing with problem-solving. In this work, we are interested in understanding how the subject is being conducted in Software Engineering-related research and more specifically, in Agile Methodologies. For that we conducted a Systematic Mapping, so we can have an overview of the research area through classification and counting contributions related to the subject. The outcome is an inventory of papers on how Diversity is being discussed in Software Engineering and Agile Methodologies. We list the most common publication venues, the frequency publication through the years, the main areas in software engineering that are interested in the subject and finally, a first discussion on how Agile Methodologies and their intrinsic characteristics can support better deliveries from more diverse development teams.**

***Index Terms*—Software Engineering, Software Development, Agile, Diversity, Heterogeneous Teams**

I. Introduction

Diversity is being discussed intensively by different knowledge areas of society and these discussions, specifically in Software Engineering, are increasing as well. More than ever, software development is a collaborative task, and different people form software development teams. Lately, it is being discussed that we have underrepresented groups, like gender, ethnic, cultural, etc. There is a lack of representativeness and unconscious bias when we talk about characteristics as ethnicity and gender, to mention a few. Technology companies must support diversity in software development teams, and this is a challenge being embraced by several companies [1][2][3]. Indeed, the work to bring the underrepresented classes to software development is of undeniable value and need. Studies say that diversity builds better teams and delivers better results, among other benefits.

Page [1] says that we cannot tell whether diversity is good or bad unless we first know what diversity is. We could see Cognitive Diversity as the differences in how we interpret, reason and solve, how we think. Identity Diversity is determined by affiliation with a social group as gender, culture, ethnicity, religion, sexual orientation, etc. [1]. Cognitive diversity is linked to better outcomes, and studies show that cognitive diversity is influenced by identity diversity (e.g., gender, race, etc.) [1]. Page [2] also says that our identities influence how we construct our lives and how others treat us; we would expect identity-diverse groups to be more cognitively diverse than homogeneous groups. Identity Diversity leads to Cognitive Diversity mainly in problem-solving and prediction.

In Software Engineering, Agile is the ability to create and respond to change to succeed in an uncertain and turbulent environment. Agile Software Development is an umbrella term for a set of methods and practices based on the values and principles expressed in the Agile Manifesto. Solutions evolve through collaboration between self-organizing, cross-functional teams utilizing the appropriate practices for their context. Agile methodologies emphasized close collaboration between the development team and business stakeholders; frequent delivery of business value, tight, self-organizing teams; and smart ways to craft, confirm, and deliver code [3].

The outcome of this paper is a systematic mapping with an inventory of papers on how diversity is being discussed in Software Engineering and Agile Methodologies. We identified the main areas that are collaborating to the subject and the specific topics being considered. This mapping is the first step of a broader project that aims to understand how agile teams are affected by diversity. Our next steps are to identify if there are benefits putting together the values from agile methodologies and cognitive/identity diverse teams and if there are, determine which ones.

The remainder of the paper is structured as follows: Section

[1] https://diversity.google/
[2] https://www.microsoft.com/en-us/diversity
[3] https://www.facebook.com/careers/diversity-report

2574-1837/19/$31.00 ©2019 IEEE
DOI 10.1109/CHASE.2019.00010

II explains the research method used. Section III presents the results of the systematic mapping about diversity in software engineering and agile methodologies. Section IV presents additional discussion on the subject, and Section V concludes the paper.

II. Research Method

The first phase of the study was to draw a systematic mapping, in which the guidelines on how to conduct a systematic review was considered along with guidelines provided for performing a systematic mapping by Petersen et al. [4].

A. Research questions

The goal of this systematic mapping study is to determine how diversity is being considered in software engineering more specifically when using Agile Methodologies. It leads to the following research questions (RQs):

- *RQ.1. What is reported in Software Engineering literature about Diversity?*
- *RQ.2. What is reported in Software Engineering literature about Diversity in Agile Methodologies?*

B. Search Protocol

The search protocol, which has been first developed by the first and second author and later been reviewed by one other researcher in software engineering, is presented summarized. A detailed version of the protocol is available online[4]. After defining a proper scope, which was initially set to cover Software Engineering and Diversity the search string were formulated. The final string is:

("Software engineer" OR "Software engineers" OR "Software Engineering" OR "software development" OR "software developer") AND (diversity OR gender OR LGBTQ OR race OR ethnicity OR "heterogeneous team")

Furthermore, some limitations were applied to the searches. The publication year was set to be between 2001, and 2018 and inclusion and exclusion criteria were defined. We also generated a list of control papers to ensure the reliability and relevancy of the searches and to evaluate the search strings: Filippova [5], Vasilescu [6], and Vasilescu [7].

III. Results

The result selection began from 3709 papers returned by the databases after running the search string. After applying the inclusion and exclusion criteria over the keywords and the titles, the number of papers diminished to 370. In the second phase, to eliminate the duplicated papers we grouped the results from the three databases. Also, we read the abstracts to understand if more papers could be removed. After duplicated removal and abstracts reading, we got 221 papers [5]. The last phase was the full read the remaining papers. Results are shown in Table I.

[4]http://karinakohl.github.io/publications/CHASE2019/ProtocolSysMap.pdf
[5]http://karinakohl.github.io/publications/CHASE2019/BibSysMap.pdf

TABLE I
Inclusion/Exclusion Criteria

Database	# of papers	First Filter [a]	Final [b]
IEEE	957	116	
ACM	1203	94	
SCOPUS	1549	160	
Total	3709	370	221

[a] Keywords and Titles.
[b] Duplicated removal and Abstracts Reading.

A. Frequency of Publication

The outcome of the final filter was 221 papers. Fig. 1 shows the number of papers for each studied year from 2001 to Dec 2018. From 2001 to 2013, 13 years, we sum up 102 papers (no publication found in 2001). From 2014 to 2018, five years, we found 119 papers. The distribution indicates the interest in the subject increased during the last five years especially in 2018 when the number of paper about the subject reach the highest: 37 papers. A highlight for 2018 is the occurrence of the First Workshop on Gender Equality in Software Engineering, contributing with ten papers in this systematic mapping and pushing the numbers. In 2013, we can observe a decrease in the number of papers to just after an increase. A hypothesis is because, in the last few years, society began to talk intensively about diversity, mainly in STEM (Science, Technology, Engineering, Mathematics) areas. Companies like Facebook[3] and Google[1] began to publish their annual diversity reports in 2014, and Microsoft[2] from 2016 on. Marissa Mayer took the CEO position at Yahoo! in 2012, and Sheryl Sandberg the COO position at Facebook in 2012 and she also has a seat in the board of directors (the first woman). In 2015, a post from Sherman [8] mentioned that in 2013, only 26% of computing professionals were female - down considerably from 35% in 1990 and virtually the same as in 1960. There are pieces of evidence that the topic returns to conversations around the period.

B. Venues

In this study, we considered peer-reviewed venues, including books, journals, conferences and workshops. We have 146 publications about the subject in conferences, 57 in journals, 15 in workshops, and three in books.

We also identified the venues were mostly targeted by authors on the subject. The top venues that published about Diversity in Software Engineering and Agile Methodologies are: with ten publications, FIE (Frontiers in Education), GE (Workshop on Gender Equality in Software Engineering (ICSE)), and ICGSE (ACM/IEEE International Conference on Global Software Engineering). ICSE (ACM/IEEE International Conference on Software Engineering) with eight publications, VL/HCC (IEEE Symposium on Visual Languages and Human-Centric Computing), seven publications, and IEEE Software with six publications. The conferences are important conferences in Software Engineering and the journal *IEEE Software* is a peer-reviewed magazine and scientific journal published by the IEEE Computer Society covering all aspects of software

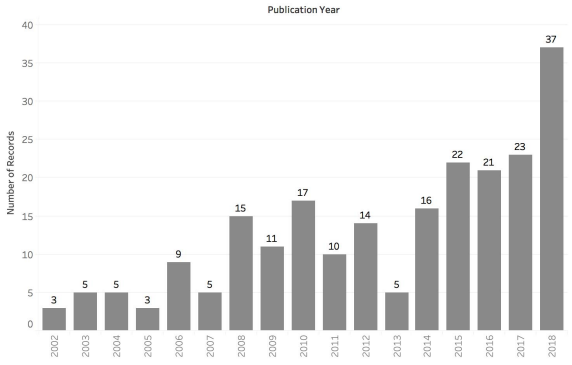

Fig. 1. Number of Papers per Year

engineering, processes, and practices. It indicates that the studies are regarded as valuable scientific contributions, given that they are published in high-quality forums.

C. Research Questions Answers

We found out 209 papers that answer our RQ.1 and 12 papers related to Agile Methodologies that help us answer the RQ.2. Papers from [P1] to [P209] answer the RQ.1, and papers from [P210] to [P221] answer RQ.2. Two of the papers that respond RQ.2 overlaps with RQ.1 in GSD/GSE topic, and one overlaps Others(Teams Management) topic.

1) RQ.1: From the relevant research, all the papers provide answers to the RQ.1 but in different topics regarding Software Engineering, as Global Software Development/Global Software Engineering (GSD/GSE) with 31 publications, Education with 65 publications, CSCW (OSS/Social) with 37 publications, and Other (Human-Computer Interaction, Teams Management, Quality, Research) with 99 publications. Some topics may overlap, and besides the mixed content under "Other" categories, the topic that is highly interested in diversity is Education.

2) RQ.2: From the twelve papers that answer RQ.2, seven are studies that relate gender diversity and agile methodologies straightforward. Like the papers that answer RQ.1, most of the papers that answer RQ.2 regarding diversity are related to gender and women. One paper mentions LGBTQI (lesbian, gay, bisexual, transgender, queer, intersex) briefly [P212]. Three papers use the Scrum framework as an enabler to have more efficient, diverse teams [P213][P218][P220]. One talks about the impact of having a female scrum master, proposing that the team can benefit from female characteristics in this role [P217]. Four papers address how Agile Methodologies can support better performances in teams and how agile values can help to create psychological safety in software development environments, even helping to retain women in the role of software developers [P214][P215][P219][P221]. Three papers are about pair programming. One is about how pair programming and agile can foster collaborative learning environments [P210]. The other two are studies about different gender pair configurations (Female-Female, Female-Male, and Male-Male) and the output of these pairing. Both ran quantitative and qualitative experiments in academic environments and found out that in quantitative terms, there is no difference in how much code is delivered (productivity is the same).

IV. Discussion

All the 221 papers answer our RQ.1, and a subset of 12 papers related to Agile Methodologies help us to answer the RQ.2. We are also interested in understanding what kind of diversity the paper is approaching, identity or cognitive. The characteristics used to categorized between identity or cognitive diversity were extracted from the words/keywords used in the papers and then put it in the equivalent diversity definition based on the literature used in this paper (Page [1]). Table II shows the distribution. It may have overlap in both categories, and some papers address only identity diversity.

TABLE II
Number of Papers About Identity and Cognitive Diversity

Identity		**Cognitive**	
Gender	129	Ethics	2
LGBTQI	2	Personality	17
Age	10	Functional	4
Race	7	Psychological	7
Cultural	67	Thinking	7
Disabilities	1	Background	24

In the identity diversity perspective, from the 221 qualified papers, 129 address gender diversity, and a subset of 83 directly mentions women. Two papers address LGBTQI, and one overlaps with gender topic. Ten papers approach age/generation diversity, overlapping four with gender, one with race/ethnicity, and four with cultural diversity. Seven papers talk about race/ethnicity, and one overlaps with age, four with gender, and two with culture. Only one paper deal with disabilities and 67 papers deal with cultural diversity. When we restrict to the papers related to agile methodologies, seven papers addresses gender, a proportion similar to the entire set of qualified papers (58%).

From the cognitive diversity perspective, we do not have the 221 papers classified because most of the papers address identity diversity as the main topic. We analyzed the papers that talked about cognitive diversity and if there are relations with identity diversity. Two papers approach ethics overlapping with the cultural diversity from the identity diversity point of view. Seventeen papers talk about personality diversity and link with five papers about gender diversity. Four papers address functional diversity overlapping one with gender and one with cultural diversity. Seven papers approach Psychological diversity, with three overlaps with gender, two with age, and one with cultural diversity. Seven papers mentioned the diversity of thinking, three overlapping with gender and two with cultural diversity. Last, 24 papers talks about the diversity of background and nine intersect with gender, two with age, and eighth with cultural diversity. We have 39 papers about diversity that create the link between identity diversity and cognitive diversity. As mentioned earlier, Page

[1] says that identity diversity leads to cognitive diversity. He also demonstrates how collections of individuals with diverse tools can outperform collections of high "ability" individuals at problem solving and predictive tasks. In problem-solving, these benefits come from combinations of tools that can be more powerful than the tools themselves. In predictive tasks, diversity in predictive models reduces collective error [9].

From the Agile point of view, in topics as gender, age, race/ethnicity, and cultural diversity, there is some evidence of a relation with cognitive diversity in four of the twelve papers, linking to functional, psychological, thinking and background diversity. Considering that problem-solving and predictive tasks are usual regarding software engineering and agile methodologies, it seems there is an opportunity to understand better how identity diversity leads to cognitive diversity in these areas and expand the studies.

Most of the papers address only gender, specifically, from women point of view. It sounds reasonable to expand researches to include different approaches to diversity. An example of that is McKinsey Company. In 2007 they published their first report called "Women Matter" [10] identifying a positive relationship between corporate performance and high presence of women in the workplace in several Western European countries. In 2015, they published "Diversity Matters," [11] expanding the focus of their research on the relationship between performance and diversity to include diversity in race and ethnicity and sexual orientation as well as gender. They identified that companies in the top quartile for gender diversity were 15% more likely to have financial returns that were above their national industry median, and the companies in the top quartile for racial/ethnic diversity were 35% more likely to have financial returns above their national industry median. They mentioned that the correlation does not prove that the relationship is causal - that greater gender and ethnic diversity in corporate leadership automatically translates into more profitbut rather indicates that companies that commit to diverse leadership are more successful. The existence of the relationship is statistically significant and consistently present in the data.

Those insights corroborate on what we are expecting as next steps, mentioned earlier in this systematic mapping: to identify if there are benefits in agile diverse teams, to determine which ones, and to understand the best software development team configuration regarding diversity aspects to have better performance and results.

V. Conclusion

We presented a systematic mapping of the literature from 2001 to 2018 about Diversity in Software Engineering and Agile Methodologies. We identified a list of 221 qualified papers and evaluated them concerning how Diversity pervades Software Engineering and Agile Methodologies. Research on diversity aspects of Software Engineering is being done for some time, but when regarding Agile Methodologies, specifically, only 12 studies were found.

As mentioned earlier in this paper, this systematic mapping is the first step of a broader project that aims to understand how agile teams are affected by diversity. There are research studies about Diversity in Software Engineering, but the literature is missing papers by how Diversity impacts Agile Methodologies. Future research in this field looks like a promising area of study. Once the subject that diversity leads us to better teams and better results is widely discussed, our next steps are to identify which are the benefits on putting together the values from agile methodologies and diversity and how the impacts on development teams' performance. Also, there is significant space to broaden the discussion of diversity in software engineering beyond gender issues or create more synergy between gender and other diverse characteristics as race/ethnicity, age, and culture.

Acknowledgement

Este trabalho foi parcialmente financiado pela FAPERGS, atraves do projeto 17/2551-0001/205-4. This project is partially funded by FAPERGS, project 17/2551-0001/205-4.

References

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

[2] ——, *The Diversity Bonus: How Great Teams Pay Off in the Knowledge Economy (Our Compelling Interests)*. xyz: Princeton University Press, 2017.

[3] Agile alliance. [Online]. Available: https://www.agilealliance.org/agile101/

[4] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering," in *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering*, ser. EASE'08. Swindon, UK: BCS Learning & Development Ltd., 2008, pp. 68–77. [Online]. Available: http://dl.acm.org/citation.cfm?id=2227115.2227123

[5] A. Filippova, E. Trainer, and J. D. Herbsleb, "From diversity by numbers to diversity as process: Supporting inclusiveness in software development teams with brainstorming," in *2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE)*, May 2017, pp. 152–163.

[6] B. Vasilescu, V. Filkov, and A. Serebrenik, "Perceptions of diversity on git hub: A user survey," in *2015 IEEE/ACM 8th International Workshop on Cooperative and Human Aspects of Software Engineering*, May 2015, pp. 50–56.

[7] 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*, ser. CHI '15. New York, NY, USA: ACM, 2015, pp. 3789–3798. [Online]. Available: http://doi.acm.org/10.1145/2702123.2702549

[8] E. Sherman. (2015) Google diversity. [Online]. Available: http://fortune.com/2015/03/26/report-the-number-of-women-entering-computing-took-a-nosedive/

[9] S. E. Page, "Making the difference: Applying a logic of diversity," *Academy of Management Perspectives*, vol. 21, no. 4, pp. 6–20, 2007. [Online]. Available: https://doi.org/10.5465/amp.2007.27895335

[10] G. Desvaux, S. Devillard-Hoellinger, and P. Baumgarten. (2007) Women matter - mckinsey-company. [Online]. Available: https://www.mckinsey.com/ /media/McKinsey/Business%20Functions/Organization/Our%20Insights /Gender%20diversity%20a%20corporate%20performance%20driver /Gender%20diversity%20a%20corporate%20performance%20driver.ashx

[11] V. Hunt, D. Layton, and S. Prince. (2007) Diversity matters - mckinseycompany. [Online]. Available: https://assets.mckinsey.com/ /media/857F440109AA4D13A54D9C496D86ED58.ashx