The Journal of Systems and Software 217 (2024) 112162

Contents lists available at ScienceDirect

The Journal of Systems & Software

journal homepage: www.elsevier.com/locate/jss

Uncovering gender gap in academia: A comprehensive analysis within the software engineering community☆

Andrea D'Angelo *, Giordano d'Aloisio, Francesca Marzi, Antinisca Di Marco, Giovanni Stilo

Department of Information Engineering and Information Sciences and Mathematics, University of L'Aquila, Via Vetoio, L'Aquila, 67100, Italy

ARTICLE INFO

Keywords:
Gender gap
Gender bias
Academia
Italy
Informatics
Software engineering

ABSTRACT

Gender gap in education has gained considerable attention in recent years, as it carries profound implications for the academic community. However, while the problem has been tackled from a student perspective, research is still lacking from an academic point of view. In this work, our main objective is to address this unexplored area by shedding light on the intricate dynamics of gender gap within the Software Engineering (SE) community. To this aim, we first review how the problem of gender gap in the SE community and in academia has been addressed by the literature so far. Results show that men in SE build more tightly-knit clusters but less global co-authorship relations than women, but the networks do not exhibit homophily. Concerning academic promotions, the Software Engineering community presents a higher bias in promotions to Associate Professors and a smaller bias in promotions to Full Professors than the overall Informatics community.

Terminology

Following the well-established research practices (Abramo et al., 2016), in this work, we adhere to the convention where the term "gender" refers to the birth sex. We want to note that the literature commonly uses the term "gender" as a biological sex, even though we acknowledge that this might differ from its meaning regarding gender identity.

Moreover, we distinguish between the terms "gender bias" and "gender gap". We use "gender bias" to refer to discrepancies that are formally measured through the formal bias metric. Conversely, "gender gap" is used to describe behavioral mismatches or observable differences in the networking habits of men and women. This distinction allows us to discuss both the potential biases that could be influencing professional networking as well as the empirical differences in behavior.

Finally, we use the word "Researcher" to identify the Italian academic position corresponding to the international "Assistant professor".

1. Introduction

The problem of *gender gap* has been widely considered and analyzed in the literature under several contexts and domains, like health (Ruiz and Verbrugge, 1997), justice (Angwin et al., 2016), or education (Moss-Racusin et al., 2012). Concerning education, the problem of gender gap gained considerable relevance over the years, and several papers studied this issue from both a technical and sociological point of view (Baker and Hawn, 2021; Mengel et al., 2019). The relevance of this topic is also highlighted by international programs such as the 2030 United Nations Sustainability Development Goals (Nations, 2015) or European Union programs such as EUGAIN, which aims to improve gender balance in Informatics at all levels (Anon, 2020). Exploring this issue from an academic perspective can provide valuable insights into the systemic challenges that perpetuate the gender gap within educational institutions, enabling the development of evidence-based strategies and policies to promote equity and inclusion. However, most works focus on the gender gap in student's education, not including other relevant contexts (Baker, 2023). In this work, we tackle the issue of the gender gap in education from an academic and multi-faceted point of view, focusing on academic promotions in Software Engineering (SE).

To achieve this goal, we collect all the needed data from several open repositories and process them with data mining techniques to make them suitable for analysis. The collected dataset, reusable for further evaluations, is thoroughly anonymized. We conduct a comprehensive analysis with a dual focus, comparing the Italian SE community with the worldwide SE community and the Italian Informatics community, analyzing both co-authorship relationships and the scenarios of academic promotions.

☆ Editor: Professor Laurence Duchien.

* Corresponding author.
E-mail addresses: andrea.dangelo6@graduate.univaq.it (A. D'Angelo), giordano.daloisio@graduate.univaq.it (G. d'Aloisio), francesca.marzi@univaq.it (F. Marzi), antinisca.dimarco@univaq.it (A. Di Marco), giovanni.stilo@univaq.it (G. Stilo).

https://doi.org/10.1016/j.jss.2024.112162
Received 8 November 2023; Received in revised form 24 May 2024; Accepted 20 July 2024
Available online 23 July 2024
0164-1212/© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

The dataset of reviewed papers, together with the anonymized dataset we collected, and the necessary code to replicate the experimental evaluation are available on Zenodo (D'Angelo and d'Aloisio, 2024).

In particular, in order to spotlight the current situation about the gender gap in SE and Informatics communities, we aim to answer the following research questions:

RQ1. Do researchers of different gender groups differ in their behavior when selecting co-authors, including whether they consistently work with the same group or not?
This research question aims at highlighting involuntary behavioral patterns that hinder cooperation between researchers of different gender groups. As women are underrepresented in SE, any tendency for same-gender collaboration will disproportionately affect them and reduce their visibility, which is critical for career advancements. To address this question, we build social graphs based on co-authorship relations between authors in the worldwide SE and Italian SE Communities. We perform a Network Analysis on these graphs and evaluate several metrics, including Homophily, Clustering Coefficient, and Modularity.

RQ2. How is the Italian software engineering community dealing with gender bias in academic promotions with respect to the overall Italian informatics community?
This RQ aims at analyzing the gender bias in academic promotions in the SE and informatics communities. Performing an analysis of specific academic areas or countries is relevant because each of them has a distinct set of implicit and explicit promotion rules (Schimanski and Alperin, 2018). Tackling this question needs additional data on the career trajectory of the researchers involved, which is challenging to collect on a global scale since different countries have different criteria for promotion. This is the reason why we restrict our analysis to the Italian context for this question. We gather the necessary public data from the Italian minister of education (*"Ministero dell'Istruzione e del Merito - MIUR"*) (dell'Istruzione dell'Università e della Ricerca, 2023), and we employ Disparate Impact (DI) (Feldman et al., 2015), a formal bias metric, to evaluate and compare the gender bias in academic promotions in the Italian SE community and the Italian Informatics community.

RQ3. What is the degree of gender bias in academic promotions between people with a high number and people with a mid-low number of publications and citations in the Italian informatics and SE communities?
The Italian academic promotion system in STEM is strongly based on bibliometric indexes like the number of publications and the number of citations (Anon, 2024a,b). This RQ takes into account this factor by assessing if, given the same academic performance (i.e., number of publications and citations), there is still a gender bias in academic promotions within the Italian informatics and software engineering communities. To answer this question, we perform a deeper analysis by employing the number of citations and publications for each author. In particular, we divide the SE and informatics populations into two sub-groups: people with a number of publications and citations greater or equal to the 75% of the whole population (called *Best Performers*), and people with a number of publications and citations lower than the 75% of the population (called *Mid-Low Performers*). Next, we compute the yearly DI for Best Performers and Mid-Low Performers groups and compare the results between the SE and informatics communities.

This paper is an extension of d'Aloisio et al. (2023), where we investigated only RQ2.[1] In this work, we strengthen our study by:

- considering more than double the papers within the current literature on the subject of gender gap in SE and academia.
- expanding our data collection to the Worldwide SE community, building co-authorship graphs, and performing Network Analysis to unearth common patterns. (RQ1)
- expanding the evaluation of DI on the communities' Best performers and Mid-low performers, whereas in our previous work, we only considered the overall community. (RQ3)

This dual assessment of co-authorship networks and academic promotions provides a comprehensive view of gender gap within the SE field. It offers insights into the collaboration dynamics among researchers and the biases affecting their career advancements, allowing us to infer significant results on the state of gender gap within the worldwide SE, Italian SE, and Italian Informatics Communities. Moreover, as demonstrated by other works on the field of gender equality in SE (Tahsin et al., 2022, 2023; Motogna et al., 2022), highlighting current challenges and insights about gender equality in the informatics and SE academic communities can increase the awareness on such topics and motivate further research on the field.

This paper is structured as follows. In Section 2, we present the study on the existing literature on gender gap in SE and academia. In Section 3, we detail the full methodology employed for the Network Analysis and quantitative bias evaluation. In Section 4, we present the results of the evaluation of the Network Analysis and Disparate Impact computation. Section 5 presents threats of validity to this study following the classification scheme reported in Wohlin et al. (2012). Lastly, in Section 6, we delineate the conclusions of our work and possible future work.

2. Related work

In this section, we give an overview of how the issue of gender gap has been analyzed both in the SE community and in the academic context. To this aim, we select related works from both relevant venues on SE or, more generally, addressing the gender gap in academia.

2.1. Gender gap in software engineering

To address how the issue of gender bias in the SE community has been analyzed over the years, we review papers from the following relevant venues about gender equality in SE: "Gender Equality, Diversity, and Inclusion in Software Engineering" (GE@ICSE) workshop, the "Software Engineering in Society" track of the ICSE conference, and the EUGAIN European project. To avoid bias in the selection of papers, following the process described in Wohlin et al. (2012), two authors independently read the title and abstract of all the accepted papers and selected papers specifically focusing on gender discrimination and inclusion. Papers selected by both authors were included, while papers selected by only one author were discussed together and eventually included if relevant.

Several works analyze the issue of gender gap in SE from an industry perspective, either addressing the current status of women's employment in SE fields (Wang et al., 2023; Qiu et al., 2023) or providing suggestions and insights on how to increase the participation of women in STEM fields (Trinkenreich et al., 2022; Tahsin et al., 2022; Kovaleva et al., 2022b,a; Zhao and Young, 2023; Szlavi et al., 2023). Concerning works analyzing the gender gap issue from a SE academic perspective, Santiesteban et al. investigate the issue of gender discrimination in the teaching evaluation of students on SE courses. Through an analysis of private data from an American university, the authors show how there is an intrinsic gender bias in the evaluation of professors (Santiesteban et al., 2022). Tahsin et al. discuss instead how an awareness-based approach can help address the gender gap issue in Bangladesh universities (Tahsin et al., 2022). Following a similar awareness-based approach, Hyrynsalmi reports the diversity and inclusion approaches

[1] We kept RQ2 because the results are linked with the new RQ3.

adopted by several SE teachers in Finnish universities (Hyrynsalmi, 2023). Marquardt et al. analyze instead how single-gender interdisciplinary classes can help women's engagement in SE fields (Marquardt et al., 2023). Similar to the work proposed in this paper, Nabot et al. analyze the trend of women's participation in computer science studies in Spain (Nebot and Mugica, 2023). Finally, Motogna et al. perform a qualitative analysis of the issues impacting women's academic careers in computer science in the context of Romania (Motogna et al., 2022).

From this analysis, we observe how there is a strong interest in analyzing the gender gap issue in the SE community. Several works have been proposed to analyze and enhance women's participation both in SE industries and academia. However, no papers study the gender gap issue in academic promotions in SE and informatics communities. This work aims to fill this unexplored area and increase awareness of current challenges and insights about gender equality in academic positions by analyzing data to understand, through a data-driven systematic approach, whether there are mechanisms, patterns, or dynamics that favor the gender gap.

2.2. Gender gap in academia

Differently from the previous section, we make a wider analysis considering papers dealing with the gender gap in other academic domains. Our aim is to identify related work on the gender gap to identify approaches similar to ours to consider for wider consideration. In fact, our aim is also to analyze how the issue of the gender gap in academia has been analyzed over the years. To this goal, we performed a light-weight systematic mapping study to select relevant works. In the following, we summarize the outcome of our study, while the detailed methodology and the full results are reported in Appendix A.

Analyzing the issue of the gender gap in academia has gained considerable relevance in the literature over the last years, as also motivated by actions from the European Union (EU) like the EUGAIN project (Anon, 2020) or by the United Nations (ONU) 2030 Sustainment Development Goals (like *Goal 4: Quality Education* or *Goal 5: Gender Equality*) (Nations, 2015). These actions from ONU and EU can also explain how the majority of works focus on academic contexts either in Europe or in North America (i.e., the western side of the world). Concerning single countries, we want to highlight how Italy is the country with the most papers on this topic after the US. The reason for this trend can be explained by the fact that the MIUR released public data about people employed in the academia (dell'Istruzione dell'Università e della Ricerca, 2023).

However, most of the proposed analyses are related to broad academic areas , and, in particular, currently, there is no paper addressing the issue of gender bias within the software engineering and informatics communities. Instead, analyses focusing on a specific academic area are needed since each area has a typical set of implicit and explicit promotion rules (Schimanski and Alperin, 2018). Moreover, we notice how most papers use data coming from private sources (like surveys or admission letters) , making such analyses difficult to reproduce and extend. In this paper, we analyze the issue of gender bias in academic promotions in Italian software engineering and informatics communities. Following the trend of other papers focusing on the Italian context, we also use data from the MIUR public database, but, differently from other works, we integrate these data with other international sources like Scopus or Google Scholar (refer to Section 3.1 for a more detailed description of the data collection process). In addition, we employ both a descriptive and inferential statistical approach using network analysis techniques and a formal bias metric (the Disparate Impact (Feldman et al., 2015)) to assess the amount of gender bias in academic promotions.

3. Experimental settings

In the following, we describe the data collection pipeline, methodology and experimental settings for the Network analysis 3.1.1 and the evaluation of Disparate Impact on the collected data 3.1.2. A comprehensive description of the collected data is then presented in Section 3.2. Fig. 1 depicts the full experimental settings workflow. The data sources we employ are both Italian and international. As international sources, we use data from the Scopus API (Elsevier, 2023) and from Google Scholar (Scholar, 2023). Concerning Italian sources, we extract data from the MIUR website (dell'Istruzione dell'Università e della Ricerca, 2023), which contains all the information about people employed in Italian academia. Starting from the collected data, we employ two different processing pipelines. The pipeline for Worldwide/Italian SE (shown on the left side of Fig. 1 and described in detail in Section 3.1.1) leverages data from Scopus and Google Scholar to obtain co-authorship graphs of both the Worldwide and Italian SE Communities. These graphs are subjects to our Network Analysis and Academic Performance evaluation, aimed to answer the **RQ1.** In particular, we extract patterns exhibited by both gender groups when choosing co-authors and publishing their work. The second pipeline (described in Section 3.1.2 and shown on the right side of Fig. 1) integrates the international data extracted from Scopus and Google Scholar with the career data obtained from the MIUR website. The collected data are then processed and filtered (using the process described in Section 3.1.2) into two distinct datasets: one having information on the Italian SE community (*SE Datasets* in Fig. 1) and one having information on the overall Italian informatics community (*INF Datasets* in Fig. 1). We use these datasets to evaluate the Disparate Impact (Feldman et al., 2015) (i.e., the formal bias metric of choice) on the promotion patterns of Researchers to Associate Professors and Associate Professors to Full Professors on the overall INF and SE Italian communities to answer **RQ2.** It is important to note that the academic role labeled as "Researcher" in Italy corresponds to what is typically known as "Assistant Professor" in other countries. Finally, the SE and INF datasets are further filtered to identify people with a number of publications and citations higher than 75% of the whole population (i.e., *Best Performers*). We then compute the DI on the Best Performers and compare the results with the DI computed on the rest of the population (i.e., *Mid-Low Performers*) to answer the **RQ3.**

3.1. Pipelines description

3.1.1. Graphs construction pipeline and selected metrics

This Section describes in detail the pipeline depicted on the left in Fig. 1 (i.e., the *Worldwide/Italian SE* pipeline). The goal of this pipeline is to build graphs for the subsequent Network Analysis and Academic Performance Evaluation. We then describe the metrics we selected for these tasks.

The data employed for this pipeline are the data we collected from Google Scholar and the Scopus API (as described in Section 3.2) to build a dataset of SE. The dataset is then further processed to construct a worldwide social graph of SE, based on the co-authorship relationships between them. In particular, each node refers to a researcher, and there is an edge between nodes if the researchers are co-authors in some of their papers. This graph is the subject of the network analysis presented in Section 4.1. From the worldwide social graph, we obtain the subgraphs of Italian SE by filtering the worldwide SE dataset using information from Google Scholar.

The Network Analysis aims to unearth patterns in the co-authorship selection among researchers of different gender groups. To this end, we selected metrics that are commonly used in the field of social network analysis to quantify the strength and nature of relationships within the co-authorship network. We perform a comprehensive Network Analysis on these graphs by measuring the following properties:

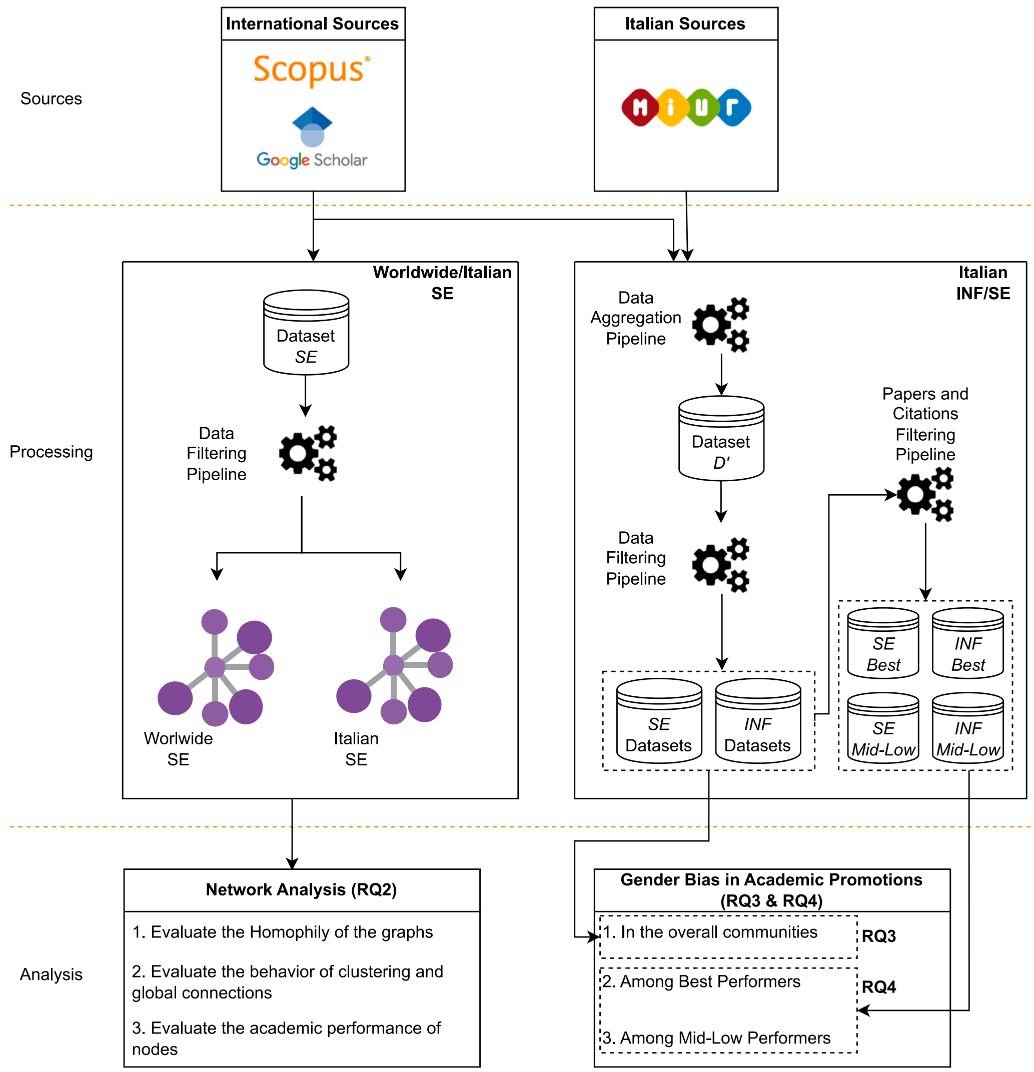

Fig. 1. Experimental Settings pipelines, divided in Sources, Processing, and Analysis.

- **Homophily**: Homophily is defined as the tendency of a group of individuals to seek connections with similar individuals. In our context, homophily is the tendency of researchers to co-author papers with researchers of the same gender. We measure homophily via two different metrics:
  - *Homophily*: defined as the ratio of cross-gender edges over all edges (Currarini et al., 2016). This is the probability that, given a randomly chosen edge in the graph, the endpoints will be of different genders. Formally, this value is

    $$H = \frac{|e_{cross}|}{|e|},$$

    where e_{cross} is the set of cross-gender edges. Assuming a ratio p of men, and a ratio q of women, the ideal value, indicating no homophily, is $w = 2pq$.
  - *Coleman Homophily Index*: normalizes the homophily with the maximal value of excess homophily. The formula is

    $$C_h = \frac{H - w}{1 - w}$$

    with H and w defined as before (Currarini et al., 2016). In this case, the ideal value is 0.
- **Clustering Coefficient**: The Clustering Coefficient measures, from 0 to 1, the tendency of nodes in a graph to cluster together in tightly-knit groups. In our context, this is the tendency of researchers to co-author papers with a fixed set of people.
  More specifically, the Clustering Coefficient of a node u is defined as Saramäki et al. (2007):

  $$CC(u) = \frac{2|T_u|}{deg(u)(deg(u) - 1)}$$

  where T_u is the set of triangles through node u, and $deg(u)$ is the degree, i.e. number of edges, of node u. A triangle through node u exists if $\exists v, m \in V | (u, v), (v, m), (m, u) \in E$, or simply put, if node u is part of a clique of size 3. The Clustering Coefficient of the entire graph that we compute in Section 4 is the mean of the Clustering Coefficients of its nodes.
- **Modularity**: Modularity quantifies network structure by computing the extent of division into clusters, reflecting the strength of connections with tightly-knit groups of nodes. In our context, this metric measures the ability of researchers to build global connections outside of their working group. Modularity is defined as (Newman, 2010):

  $$M = \frac{1}{2m} \sum_{ij} (A_{ij} \gamma \frac{deg(i)deg(j)}{2m}) \delta(c_i, c_j)$$

  where m is the number of edges, A is the adjacency matrix of the graph, deg(i) and deg(j) are the degrees of i and j, respectively, and $\delta(c_i, c_j) = 1$ if nodes i and j are part of the same community,

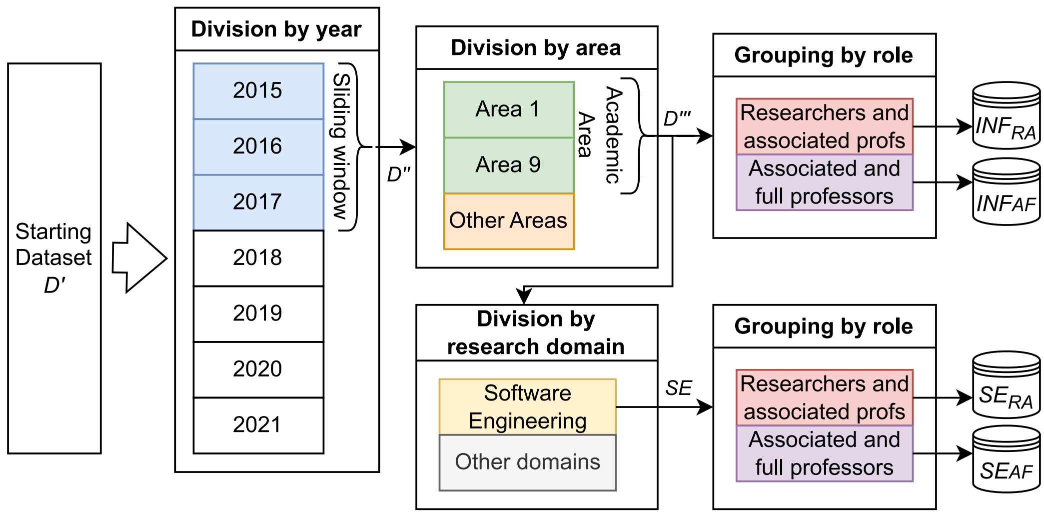

Fig. 2. Filtering pipeline of the Italian datasets.

and 0 otherwise. γ is the resolution parameter that was set to 1 for our experiments.

After the Network Analysis, we also evaluate the academic performance of the two gender groups in both the Worldwide and Italian SE Communities. To this aim, we consider the number of publications and citations obtained by Scopus for each year, from 2015 to 2022, of each researcher, classifying them by gender.

3.1.2. Italian informatics and SE pipeline and selected metrics

This section describes the pipeline shown on the right side of Fig. 1 (i.e., the *Italian INF/SE* pipeline), employed to evaluate the degree of gender bias in the academic positions within the overall informatics (INF) and software engineering (SE) Italian communities. We then describe the formal bias metric we employ to this aim. The informatics community is the conjunction of Areas 1 and 9 of the MIUR scientific areas classification dell'Istruzione dell'Università e della Ricerca (2023).

This pipeline merges the data from both Italian and international data sources into a single dataset using the *name*, *surname*, *email*, and *affiliation* as join keys.

From such dataset, we provide an anonymized dataset D' on which we perform a set of filtering operations to obtain the final datasets that we use to compute bias metrics yearly. The filtering procedure is depicted in Fig. 2. Since we are interested in the evolution of bias in academic promotions year by year, the anonymized dataset D' is split according to a sliding time window of fixed size. In particular, we considered a sliding window of three years, starting from 2015. Hence, to gather metrics for 2019, with the sliding window size set to 3, we would slice D' from 2016 to 2018. After this operation, we obtain a partially filtered dataset D'' for each sliding window.

The subsequent step selects only specific scientific areas from D''. Because different domains have different promotion criteria, it would be incorrect to group them together. This study focuses on Areas 1 and 9 of the MIUR scientific areas classification, which refers broadly to Science, technology, engineering, and mathematics (dell'Istruzione dell'Università e della Ricerca, 2023). In this study, we refer to the conjunction of these two areas as the Informatics community. By selecting only researchers in these areas, we obtain a dataset D'''. From D''', we perform two different operations. In the first, D''' is split into two versions: one without records representing researchers (INF_{AF}) and one without Full Professors (INF_{RA}). In the second phase, D''' is refined by selecting individuals working in the SE field. To this aim, we leverage Google Scholar to find individuals who have expressed interest in *software engineering* or the following related topics: *software architecture*, *model-driven engineering*, *software quality*, and *software testing*. The SE dataset is then divided into two sub-datasets as done above: one consisting of only Researchers and Associate professors (SE_{RA}), and the other consisting of only Associate and Full professors (SE_{AF}).

As a result of the data pre-processing pipeline, four distinct datasets are created. Two of them are for the overall Italian INF community (INF_{RA} and INF_{AF}), while the other two are for the Italian SE community (SE_{RA} and SE_{AF}). It is worth noting how each dataset obtained as a result from the pipeline in Fig. 2 represents the promotion to a specific role (i.e., from Researchers to Associated professors, and from Associated to Full professors). Finally, we only preserve data for workers employed at an Italian university for the entire time window. This decision was made to ensure that our analysis is not influenced by individuals achieving promotions in other countries, with possibly different criteria.

Once the final yearly datasets INF_{RA}, INF_{AF}, SE_{RA}, and SE_{AF} are constructed, the experiments can occur.

As mentioned, the experiment aims to measure the amount of gender bias in academic promotions and analyze its variation over the years. To calculate the amount of bias, we use the Disparate Impact metric (Feldman et al., 2015). This metric measures the probability of having a *positive outcome* while being in the *privileged* or *unprivileged* group and is defined formally as:

$$DI = \frac{P(Y = y_p | X = x_{unpriv})}{P(Y = y_p | X = x_{priv})}$$

where Y is the label, y_p is the positive outcome, X is the sensitive variable, and x_{unpriv} and x_{priv} are the values identifying the unprivileged and privileged groups, respectively. The more this metric is close to one, the fairer the dataset.

In our context, the label assigned to a person represents their position for that particular year. In the analysis between Researchers and Associate Professors, the positive label is *Associate Professor*, while in the analysis between Associate and Full Professors, it is *Full Professor*. The sensitive variable is *gender*, where *men* and *women* are the privileged and unprivileged groups, respectively.

The experiments we perform are threefold: on the Overall communities, on the Best performers, and on the Mid-Low performers. For the Overall communities, for each final yearly dataset (INF_{RA}, INF_{AF}, SE_{RA}, and SE_{AF}) and for each year in the considered range (2018–2022), we compute the DI between the two subsets contained in the dataset (either Researchers and Associate Professors or Associate Professors and Full Professors). We also compute the cardinality of each subset per year. Next, for each year, we compute the third

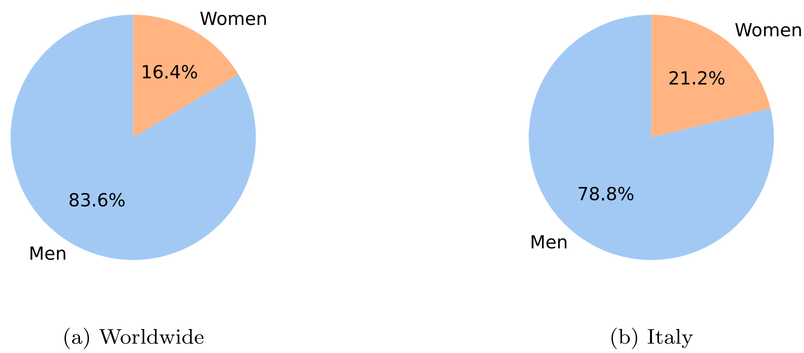

Fig. 3. Gender distribution of the Software Engineering communities, worldwide and in Italy.

quartile (75%) of the total number of publications and citations and select people having a number of publications and citations greater or equal to that value. This group is called *Best Performers*. The rest of the population is called *Mid-Low Performers*. Finally, we compute the yearly DI for both Best and Mid-low performers following the process described above along with the set cardinalities.

3.2. Data description

In this section, we give an exhaustive description of the graphs and datasets obtained at the end of the processing stage of the pipelines shown in Fig. 1.

3.2.1. Graphs description

From the Network Analysis pipeline (on the left in Fig. 1), we generate two distinct co-authorship graphs representing the Worldwide and Italian Software Engineering communities, respectively. To construct these graphs, we collected data from Scopus and Google Scholar. Specifically, we compiled a list of Software Engineers by analyzing topic tags on Google Scholar, followed by gathering detailed academic data for these researchers through the Scopus API.

In the final graphs, for each node, we store the gender, the yearly publications and citations in the last century, the publication types (books, conference proceedings, journals), and co-authorship relationships. As gender data is not available from Scopus, we employ automatic tools of gender detection given the First Name. All personal information is then promptly anonymized via ID numbers.

The graph of Worldwide Software Engineers consists of 14 661 edges and 3665 nodes, of which 3064 men and 601 women. The graph of Italian Software Engineers consists of 393 edges and 99 nodes, of which 78 are men and 21 are women.

3.2.2. Italian INF/SE dataset description

From the gender bias pipeline (on the right in Fig. 1), we obtain two datasets: one consisting of Italian Software Engineers (SE) and another covering Italian professionals in the field of Informatics (INF). Both datasets are initially unsplit; however, for analysis purposes, they are later divided into academic performance quartiles. Our descriptions pertain to the data prior to this division.

These datasets are built combining data from Google Scholar, Scopus, and MIUR. From the topic tags available in Google Scholar we discern people in the field of Software Engineering from people in general Informatics. Then, from Scopus, we collect data pertaining to academic performance. In particular, for each record, similarly to the previously described graphs, we store the yearly publications and citations in the last century, the publication types (books, conference proceedings, journals), and co-authorship relationships. Additionally, for each record we store the gender and annual updates on academic roles (Researcher, Associate Professor, and Full Professor). These data is integrated from MIUR, thus avoiding reliance on automatic gender determination methods. Again, all personal information is promptly anonymized.

The SE dataset contains a total of 117 records, of which 99 men and only 18 women. As of 2022, 53 of them were researchers, 32 were Associate professors and 32 were Full professors. On the other hand, the bigger INF dataset contains 29.568 records. Of these, 19 330 are men and 10 238 are women. As of 2022, 14 405 of these are researchers, 9672 are Associate professors and 5491 are Full Professors.

4. Experimental evaluation

The experimental evaluation is organized as follows. In Section 4.1, we compare the Italian SE community with the Worldwide SE community. We do this by analyzing their co-authorship networks regarding clusters and homophily. We then compare the mean bibliometrics performances of the Italian and Worldwide SE communities for each gender group. In Section 4.2, we perform an analysis of gender bias in academic promotions by relying on the DI metric (Feldman et al., 2015) for the Italian SE Community and the Italian Informatics Community. In particular, we conduct an in-depth analysis of gender bias in academic promotions by analyzing different groups defined by the number of publications and the number of citations. The goal of this evaluation is to paint a complete picture of how the Italian SE Community fares in addressing gender gap against both the Worldwide SE Community and closely related subjects in Italy.

4.1. Analysis of worldwide and Italian SE communities

In this section, we show and discuss the experiments performed in the worldwide and Italian Software Engineering communities. We start by showing the results of the Network Analysis carried out on the co-authorship network of Software Engineers. These experiments were performed on a dataset of 3665 Software Engineers, of which 3064 are men and 601 are women (for the worldwide community), and a dataset of 99 people, of which 78 are men and 21 are women (for the Italian community). The gender distribution is visualized in Fig. 3.

4.1.1. Network analysis

By inspecting the social graphs of Software Engineers, we can infer interesting differences between the co-authorship habits of women and men. In these graphs, each node is a researcher, and each edge is a co-authorship relation between them. In particular, in order to gather significant findings for the network behavior of the two groups, we inspect homophily, clustering coefficient, and modularity of both subgraphs of men and women. These metrics allow us to draw meaningful conclusions about the behaviors of the two groups when working together as co-authors.

Table 1 shows the homophily values according to two different metrics, where the Ideal Value means that there is no homophily. If we were to only consider the ratio of same-gender edges over all edges (which is $1 - Homophily$, 0.729) we would conclude that the

Fig. 4. Clustering coefficient of the different gender groups in the Worldwide Community (upper side) and the Italian community (lower side).

Table 1
Homophily values computed via different metrics.

Metric	*Europe*		*Italy*	
	Observed	Ideal	Observed	Ideal
Homophily	0.271	0.274	0.41	0.35
Coleman Homophily	0.099	0	−0.011	0

number of same-gender edges is significantly higher. However, it is crucial to remember that this value does not take into account the proportion of the gender groups inside the graphs. As shown in Fig. 3, this is significant to us, as the imbalance between groups is severe. Consequently, we need to compare Homophily with the ideal value $2pq$ (refer to Section 3.1.1) to account for this imbalance. As shown in Table 1, the Observed Homophily values for both Europe and Italy are very close to their ideal value, with Italy being slightly less balanced. However, this small variation might be due to a lower sample size. According to the Coleman Homophily, the Observed Value is very close to the Ideal Value of no homophily for both communities. We can thus infer that in both cases the SE community exhibits no Homophily in co-authorship relations, meaning that gender does not seem to be an important criterion for researchers when selecting co-authors for their work.

We then turn our attention to the Clustering Coefficient (Soffer and Vazquez, 2005) of the different genders in the graph. Fig. 4 presents boxplots illustrating results for the subgraphs for both men and women. Since this metric is influenced by the size of the groups, we computed the Clustering Coefficient also on a subgraph of men that has been sampled to match the number of nodes in the subgraph of women. To account for statistical variation in the sampled subgraph, we computed the mean of the clustering coefficient values for each node over 10 different samplings.

These results show that the subgraph of men exhibits a higher Clustering Coefficient with respect to the subgraph of women. While this may be partially due to the imbalance of groups, the median of the sampled subgraphs of men is still higher than the subgraph of women. This holds for both the Worldwide and Italian SE communities. We can thus infer that men generally work in tightly-knit groups more often than women. The Clustering Coefficients of men and women were also tested via the Wilcoxon–Mann–Whitney test (Fay and Proschan, 2010), which resulted in a p-value of 7.77×10^{-6}. This proves a statistically significant difference between the two groups with a 95% confidence interval. To corroborate these findings, we investigate the Modularity (Newman, 2006) of the men and women subgraphs.

We computed the normalized modularity for both subgraphs. The Modularity values are 3.5×10^{-4} and 3.3×10^{-3} for men and women in the worldwide SE community, respectively. These values are both very low and hint at the fact that neither group exhibits a strong community structure. However, the modularity of the subgraph of women is much higher in comparison. We can infer that the global structure of the subgraph of women is more pronounced and clustered with respect to the one of the men, indicating that women build global relationships outside of their working group more often than men. For Italy, the Modularity values are 0.01 for men and 0.04 for women. While these values are generally higher, the same line of reasoning can be applied to the Italian community.

4.1.2. *Metrics for academic performance*

Fig. 5 reports the mean number of publications and citations for each gender group both worldwide (upper side) and in Italy (lower side). For the former, the two gender groups have comparable academic productivity, although slightly favorable to women, who generally publish more papers and get cited more often. On the other hand, there is a stark contrast in academic productivity between the two gender groups in Italy. Women appear to publish less and get cited less than men. While part of this might be due to lower sample size and therefore higher statistical variation, the trend remains worrying and should be the subject of future research.

Lower metrics for the year 2022 are due to the data being incomplete for that year, but it is interesting to note that the mean number of papers and citations, in general, seem to be on a downtrend with respect to the peak observed in the years 2018 to 2020.

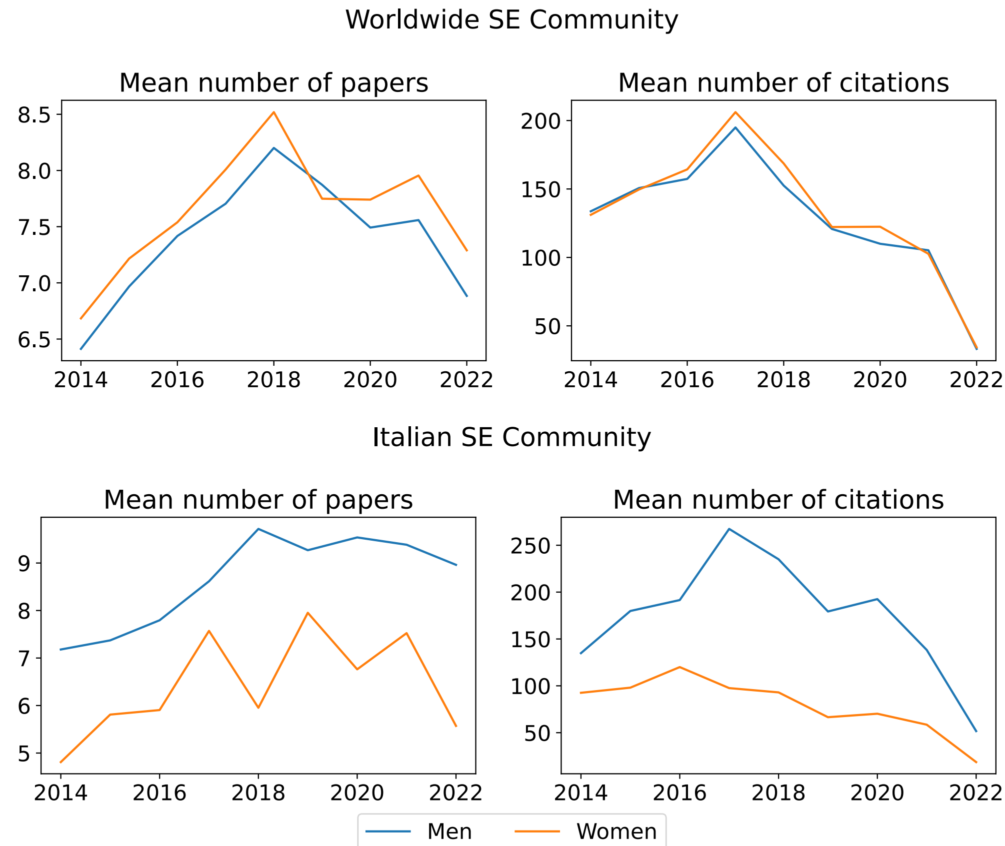

Fig. 5. Mean publications and citations in the Worldwide (upper side) and Italian (lower side) SE communities.

4.1.3. Discussion

The network analysis conducted on the co-authorship graphs pertaining to both the global Software Engineering (SE) Community and its representation in Italy conveys remarkably analogous results. Women make up, in both cases, around 20% of the SE Community (Fig. 3), and when taking this percentage into account, the networks show no evidence of homophily (Table 1), e.g., preference of choosing co-authors within the same gender group. In general, however, the subgraph of men exhibits a higher Clustering Coefficient (Fig. 4) and lower Modularity, which means that they tend to form smaller, more tightly-knit working groups with respect to women, but generally entertain less co-authorship relationships with researchers outside of their cluster. We can draw similar conclusions for both communities since the analyzed metrics are comparable and coherent between the worldwide and Italian SE communities.

The academic performance metrics of Fig. 5 show how academic performance is generally comparable between the two gender groups worldwide, but is severely different in the Italian SE Community. This is a worrying trend that needs to be further investigated.

> **Answer to RQ1:** While there is no evidence of researchers deliberately selecting colleagues of the same gender as co-authors, it is observed that men tend to work in more closely-knit clusters and have fewer connections outside of their respective clusters when compared to women. These results apply to both the Worldwide and the Italian SE (Software Engineering) communities.

4.2. Analysis of gender bias in the SE and Italian informatics communities

In this section, we present the comparison between the Italian SE community and the general Italian Informatics Community. We focus on the gender bias occurring in the events of *academic promotion*. In particular, we consider the promotions from Researcher to Associate Professor and from Associate Professor to Full Professor. The analysis is carried out on the four datasets of Informatics Researchers/Associate Professors (INF_{RA}), Informatics Associate/Full Professors (INF_{AF}), SE Researchers/Associate Professors (SE_{RA}), and SE Associate/Full Professors (SE_{AF}).

4.2.1. Disparate impact evaluation

Fig. 6 shows the DI (left y-axis) and set cardinalities (right y-axis) for each of the datasets above (INF_{RA}, INF_{AF}, SE_{RA}, and SE_{AF}) on a yearly basis in the reference period (2018–2022). In the figure, the charts on the left side show results for the Informatics (INF) Community datasets (INF_{RA}, INF_{AF}), while the ones on the right side show results for the Software Engineering (SE) Community (SE_{RA}, SE_{AF}).

Concerning the full set cardinalities (i.e., of both men and women), they exhibit the same trend across all datasets. Since we only consider people who were in the Italian academic system for the entire reference period, we do not consider Researchers who were acquired later than 2018, so their cardinality is bound to decrease. The number of Full professors is rising in both the INF and SE communities, but the increase in the SE community is significantly larger. In 2022, there are more Full professors than Associate professors specifically in the SE subset. This suggests that promotions to Full professorship are occurring at a higher rate among professors in the field of SE compared to the INF community.

Concerning the gender bias in promotions to Associate Professor (INF_{RA} and SE_{RA} in the figure), in both the Informatics and Software Engineering communities the trend of DI appears to be on an upward trajectory. However, the SE community seems to suffer from a higher bias than the overall INF community. The DI for the SE community starts from a value of 0.75 in 2018 to a value of 0.8 in 2022. In contrast,

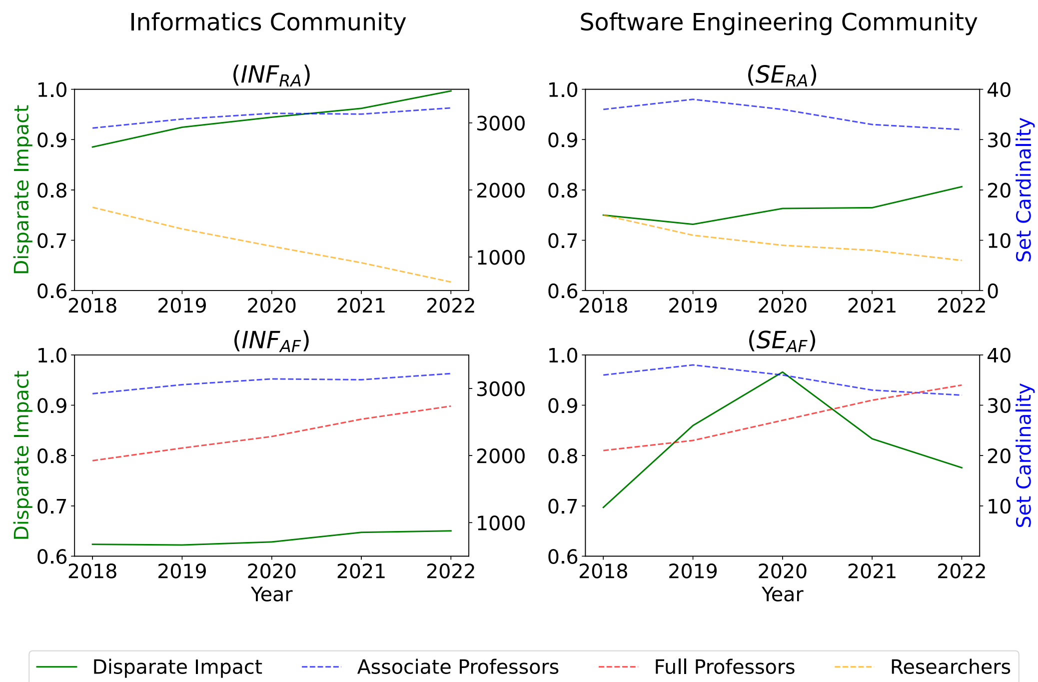

Fig. 6. Year-by-year Disparate Impact and Set Cardinality for the Informatics Community (left column) and Software Engineering Community (right column).

the DI of the INF community starts from a value of 0.9 in 2018 to a value of almost 1 in 2022, meaning a nearly complete absence of bias in academic promotions. In general, we observe how the amount of bias in the SE community is about 20% higher than in the overall INF community.

In contrast, concerning bias in promotions to Full Professors (INF_{AF} and SE_{AF} in the figure), the SE community exhibits a much lower bias concerning the INF community. DI for the SE community starts from 0.7 in 2018, then reaches a peak of 0.95 in 2020, to a final value of almost 0.8 in 2022. This downtrend from 2020 to 2022 can be partially explained by the small set cardinality, which makes the DI more sensitive to small changes (i.e., additions or deletions) in the groups. Instead, the DI for the overall INF community presents a slight increase over the period, starting from a value of 0.63 in 2018 to a value of 0.65 in 2022. In this case, the amount of bias in the INF community ranges from 15 to 35% greater than in the SE community throughout the observed period.

Answer to RQ2: The Italian SE community presents a higher gender bias in promotions from Researchers to Associate professors compared to the overall Italian informatics community. Concerning promotions from Associate to Full professors, the SE community presents a lower gender bias compared to the full informatics community.

4.2.2. *Disparate impact among best performers and mid-low performers*

We now shift our attention to the Disparate Impact exhibited by two population sub-subgroups: people having a number of publications and citations greater or equal to 75% of the entire population, and the rest of the group. We will refer to them as Best Performers and Mid-Low Performers, respectively. It is worth noticing how, as already mentioned in Section 3.1.2, people not having a Scopus ID have been excluded from this analysis, since we are not able to extract information about publications and citations for them.

Fig. 7 shows the DI (left y-axis) and set cardinalities (right y-axis) of Italian SE and informatics best performers. Is worth noticing that, in this analysis, the set cardinalities represent the total number of men and women in the specific group. In the Informatics Community, the DI when considering promotions from Researchers to Associate professors is on a downtrend, while the DI when considering promotions from Associate to Full professors is on the rise, despite a small hiccup in the most recent years. This is true despite the ratio between men and women set cardinalities remaining almost constant. This means that the probability of best performers male Researchers becoming Associate professors has increased over the years compared to the probability of female best performers Researchers. The opposite holds for Associate versus Full professors. However, it is also worth noticing that the DI concerning the promotions from Researcher to Associate professor is still very high (around 0.7 for the year 2022), meaning that, even if in a downtrend, the gender bias in such promotions is not very high. Concerning the SE community, the DI for the best performers is subject to extreme variation due to the very low number of women in this subset (with a maximum of 2 women in the comparison between Researchers and Associate professors and 1 woman in the comparison between Associate and Full professors). For certain years, the number of women with the positive label (i.e., Associate professors in case of the comparison with Researchers and Full professor in case of the comparison with Associate professors) is zero, causing the DI to be impossible to compute (we default to 0 in this case, the most biased outcome). We cannot draw meaningful conclusions in this case, but the number of women best performers in SE remains worrying and subject to future research.

Fig. 8 shows the Disparate Impact for Mid-Low Performers. In contrast with the best performers, this time the DI for both promotion scenarios in the informatics community is on an uptrend. However, while INF_{RA} reaches almost perfect fairness (DI = 1) in 2021, the level of DI for the promotions from Associate to Full professors is still low, with a maximum value of around 0.6 in 2022. From this scenario, we can conclude how, while there is almost perfect gender fairness in the promotions from Researchers to Associate professors, males still

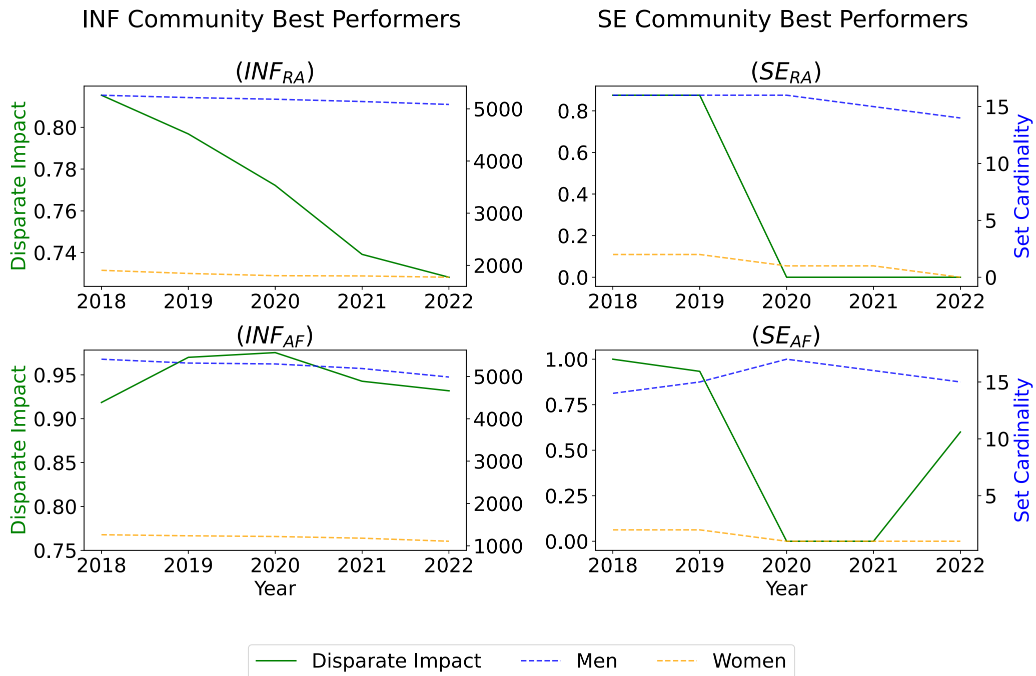

Fig. 7. Year-by-year Disparate Impact and Set Cardinality for the Informatics Community (left column) and Software Engineering Community (right column) for the Best Performers.

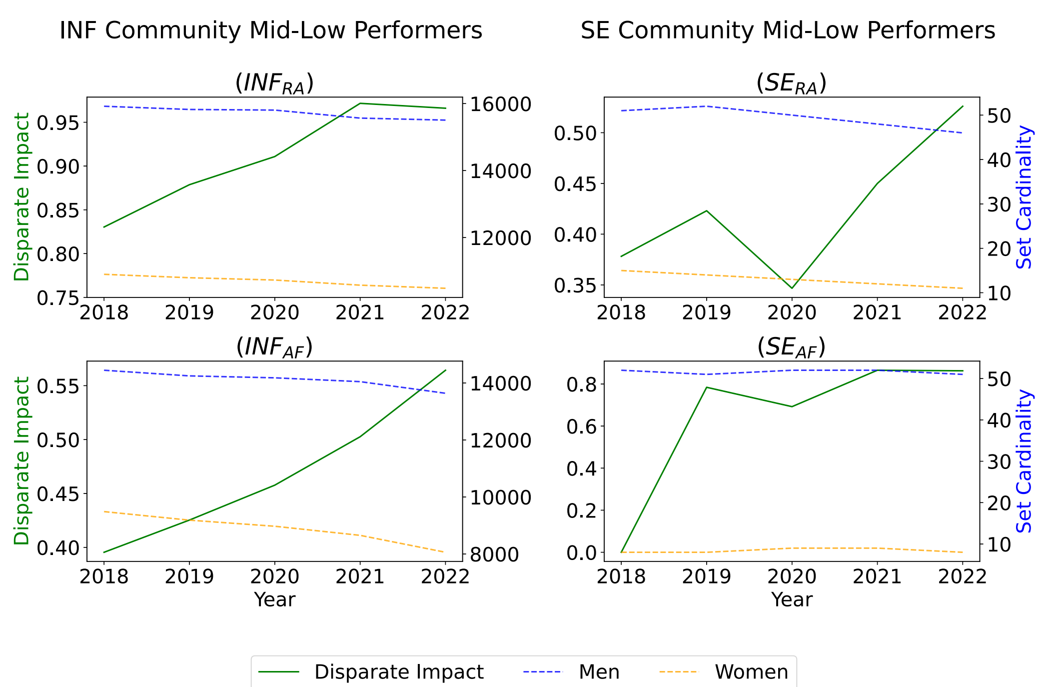

Fig. 8. Year-by-year Disparate Impact and Set Cardinality for the Informatics Community (left column) and Software Engineering Community (right column) for the Mid-Low Performers.

have a higher probability than women of becoming Full professors. On the other hand, the DI for the SE group, while still subject to heavier variation (given by the smaller dimension of the groups), is more stable with respect to the Best Performers. In particular, we observe an opposite trend with respect to the informatics community. In fact, the DI for the promotions from Researchers to Associate professors is low (with a value of around 0.5 in 2022). This means that males have a higher probability than women of becoming Associate professors. On the contrary, we observe an uptrend in gender fairness concerning promotions from Associate to Full professors (with a value of DI of around

0.75 in 2022). Finally, we note from the data how until 2019 there were no women Full professors with a mid-low number of publications and citations in the SE community.

Answer to RQ3: Concerning best performers, we observe a low gender bias in both promotions to Associate and Full professors in the Italian informatics community. Instead, the low number of women best performers in the SE Italian community does not allow us to draw meaningful conclusions about gender bias in this context. Concerning mid-low performers, we observe how there is still a gender bias in promotions to Full professors in the Italian informatics community, while we detect a low bias in promotions to Associate professors. Mid-low performers of the SE community follow an opposite trend since we observe higher gender bias in promotions to Associate professors, while there is gender fairness in promotions to Full professors.

5. Threats to validity

This section discusses possible threats that can hamper the results of the performed evaluation.

Conclusion Validity concerns issues about the results drawn from our evaluation. In this context, there may be other variables that can influence academic promotions, like the number of years in academia. To address this issue, we fix our analysis to a specific range of years and consider only people being in the academic context for the whole considered range. Moreover, we plan to extend our analysis considering also other relevant features like participation in the organizing and steering committee and in the editorial board of the most relevant conferences and journals on SE.

Internal Validity concerns internal factors that can impact the results of our evaluation. Firstly, we want to acknowledge that we are considering gender as a binary variable in our analysis despite being aware that some individuals may not identify with this classification. However, the data sources from which we have extracted gender information present it as a binary variable, limiting our ability to perform a more detailed gender analysis. Secondly, we have chosen to restrict our analysis of the Italian SE and informatics communities to individuals within the Italian academic system for all the year ranges assessed. This decision was made to ensure that our analysis of gender bias in academic promotions is accurate and not influenced by individuals leaving or joining the academic system. Lastly, we have used Google Scholar tags to identify people belonging to the SE community. While we recognize that Google Scholar tags may not be completely accurate and may not encompass all individuals in the SE community, we were unable to find more reliable sources for this information.

Construct Validity concerns how the performed experiment is adequate to address our research questions. To address this threat, we employ state-of-the-art network analysis techniques and a formal bias metric to analyze the gender bias in academic promotions. However, we publicly release the datasets used in our experiments for future experiments and analyses.

External Validity concerns the generalizability of our results. Our analysis is focused on the Italian academic system and is not applicable to countries having a different academic system.

6. Conclusion and future work

In this work, we presented an analysis of gender gap within the Italian Software Engineering and informatics communities. We started by performing a review of how the literature analyzed the subject of gender gap in academia and in the SE community. From this review, we have shown how there is growing interest in analyzing the issue of gender gap both in SE and academia. However, no papers study the issue of gender gap from a perspective of academic promotions in specific academic areas, like informatics and SE. Upon inspecting the literature, we proposed to fill the unexplored area pertaining to gender gap in the Italian Software Engineering community. We started by presenting a data collection, filtering, and processing pipeline to gather and refine academic data from multiple Italian and international public sources. We also made these datasets available online for public access. Next, we performed two sets of evaluations: (i) we compared the Italian SE community with the Worldwide SE community via Network Analysis techniques and metrics, (ii) we compared the Italian SE community with the Italian Informatics community to infer differences in the pattern of academic promotions for the different gender groups considering also differences between best performers and mid-low performers in terms of publications and citations. For the former, we infer that, despite the lack of exhibited homophily, men tend to cluster together in more tightly-knit groups and build fewer global connections with respect to women. We also conclude that the networks do not exhibit homophily, meaning that gender is not a primary factor for researchers when selecting co-authors. For the latter, we found that the Italian informatics community presents a higher gender bias in promotions to Full professors compared with the SE community. On the contrary, the SE community presents a higher bias in promotions to Associate professors. Finally, concerning the number of publications and citations, we observed a low gender bias in both promotions to Associate and Full professors for best performers in the informatics community, while the low number of women best performers in the SE community does not allow us to draw meaningful conclusions. Concerning mid-low performers, we observed a high gender bias in promotions to Full professors in the Italian informatics community, while the SE community presents a higher bias in promotions to Associate professors.

Concerning future works, an avenue of future research would be expanding the analysis on multiple countries, taking into account specific rules and regulations. Next, the Network Analysis study could be expanded upon by considering a weighted graph of co-authorship relations, which at the time of writing, were not available via the Scopus API. The Disparate Impact computation could be expanded as well by considering other factors that are relevant to academic promotions, like participation in the organization of top-tier SE conferences. Finally, while analyzing promotion patterns is important, future efforts might involve studying the degree of gender bias in the process of initial recruitment as well.

CRediT authorship contribution statement

Andrea D'Angelo: Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing, Validation. **Giordano d'Aloisio:** Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Francesca Marzi:** Methodology, Software, Writing – review & editing. **Antinisca Di Marco:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing. **Giovanni Stilo:** Conceptualization, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Link to data available on the paper.

Acknowledgments

This work is partially supported by the European Union - NextGenerationEU - National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR) - Project: "SoBigData.it - Strengthening the Italian RI for Social Mining and Big Data Analytics" - Prot. IR0000013 - Avviso n. 3264 del 28/12/2021 and by the COST Action CA19122 – European Network Balance in Informatics (EUGAIN).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jss.2024.112162.

References

Abramo, G., D'Angelo, C.A., Rosati, F., 2016. Gender bias in academic recruitment. Scientometrics 106 (1), 119–141.

Angwin, J., Larson, J., Mattu, S., Kirchner, L., 2016. Machine bias. ProPublica 23 (2016), 139–159.

Anon, 2020. CA19122-European Network For Gender Balance in Informatics (EUGAIN). URL: https://eugain.eu/.

Anon, 2024a. Italian National Scientific Abilitation. URL: https://abilitazione.mur.gov.it/public/index.php.

Anon, 2024b. Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca. URL: https://www.anvur.it/attivita/asn/asn-2012-2013/indicatori-e-relative-mediane/.

Baker, R.S., 2023. Learning analytics: An opportunity for education. XRDS: Crossroads ACM Mag. Stud. 29 (3), 18–21.

Baker, R.S., Hawn, A., 2021. Algorithmic bias in education. Int. J. Artif. Intell. Educ. http://dx.doi.org/10.1007/s40593-021-00285-9.

Currarini, S., Matheson, J., Vega-Redondo, F., 2016. A simple model of homophily in social networks. Eur. Econ. Rev. 90 (C), 18–39. http://dx.doi.org/10.1016/j.euroecorev.2016, URL: https://ideas.repec.org/a/eee/eecrev/v90y2016icp18-39.html.

d'Aloisio, G., D'Angelo, A., Marzi, F., Di Marco, D., Stilo, G., Di Marco, A., 2023. Data-driven analysis of gender fairness in the software engineering academic landscape. (accepted at ECSA 2023, soon to be published), arXiv:2309.11239.

D'Angelo, A., d'Aloisio, G., 2024. Public SE fairness. http://dx.doi.org/10.5281/zenodo.11075506.

Elsevier, 2023. Scopus. URL: https://www.scopus.com.

Fay, M.P., Proschan, M.A., 2010. Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. Stat. Surv. 4.

Feldman, M., Friedler, S.A., Moeller, J., Scheidegger, C., Venkatasubramanian, S., 2015. Certifying and removing disparate impact. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, Sydney NSW Australia, pp. 259–268. http://dx.doi.org/10.1145/2783258.2783311.

Hyrynsalmi, S.M., 2023. How diversity and inclusion are approached in software engineering university-level teaching. In: 2023 IEEE/ACM 4th Workshop on Gender Equity, Diversity, and Inclusion in Software Engineering (GEICSE). pp. 17–24. http://dx.doi.org/10.1109/GEICSE59319.2023.00007, URL: https://ieeexplore.ieee.org/abstract/document/10213434.

Kovaleva, Y., Happonen, A., Kindsiko, E., 2022a. Designing gender-neutral software engineering program. stereotypes, social pressure, and current attitudes based on recent studies. In: Proceedings of the Third Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering. In: GE@ICSE '22, Association for Computing Machinery, New York, NY, USA, pp. 43–50. http://dx.doi.org/10.1145/3524501.3527600, URL: https://dl.acm.org/doi/10.1145/3524501.3527600.

Kovaleva, Y., Happonen, A., Mbogho, A., 2022b. Towards gender balance in modern hackathons: literature-based approaches for female inclusiveness. In: Proceedings of the Third Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering. In: GE@ICSE '22, Association for Computing Machinery, New York, NY, USA, pp. 19–26. http://dx.doi.org/10.1145/3524501.3527594, URL: https://dl.acm.org/doi/10.1145/3524501.3527594.

Marquardt, K., Wagner, I., Happe, L., 2023. Engaging girls in computer science: Do single-gender interdisciplinary classes help? In: 2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET). pp. 128–140. http://dx.doi.org/10.1109/ICSE-SEET58685.2023.00019.

Mengel, F., Sauermann, J., Zölitz, U., 2019. Gender bias in teaching evaluations. J. Eur. Econom. Assoc. 17 (2), 535–566. http://dx.doi.org/10.1093/jeea/jvx057.

Moss-Racusin, C.A., Dovidio, J.F., Brescoll, V.L., Graham, M.J., Handelsman, J., 2012. Science faculty's subtle gender biases favor male students. Proc. Natl. Acad. Sci. 109 (41), 16474–16479.

Motogna, S., Alboaie, L., Todericiu, I.A., Zaharia, C., 2022. Retaining women in computer science: the good, the bad and the ugly sides. In: Proceedings of the Third Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering. In: GE@ICSE '22, Association for Computing Machinery, New York, NY, USA, pp. 35–42. http://dx.doi.org/10.1145/3524501.3527598, URL: https://dl.acm.org/doi/10.1145/3524501.3527598.

Nations, U., 2015. THE 17 GOALS | Sustainable Development. URL: https://sdgs.un.org/goals.

Nebot, A., Mugica, F., 2023. Evolution of the participation of women in university computer science studies in Spain and in europe. In: 2023 IEEE Frontiers in Education Conference. FIE, pp. 1–8. http://dx.doi.org/10.1109/FIE58773.2023.10343410.

Newman, M.E., 2006. Modularity and community structure in networks. Proc. Nat. Acad. Sci. 103 (23), 8577–8582.

Newman, M.E.J., 2010. Networks: an introduction. Oxford University Press, Oxford; New York, URL: http://www.amazon.com/Networks-An-Introduction-Mark-Newman/dp/0199206651/ref=sr_1_5?ie=UTF8&qid=1352896678&sr=8-5&keywords=complex+networks.

Qiu, H.S., Zhao, Z.H., Yu, T.K., Wang, J., Ma, A., Fang, H., Dabbish, L., Vasilescu, B., 2023. Gender representation among contributors to open-source infrastructure : An analysis of 20 package manager ecosystems. In: 2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS). (ISSN: 2832-7616) pp. 180–187. http://dx.doi.org/10.1109/ICSE-SEIS58686.2023.00025.

dell'Istruzione dell'Università e della Ricerca, M., 2023. Cerca Università. URL: http://cercauniversita.cineca.it/php5/docenti/cerca.php.

Ruiz, M.T., Verbrugge, L.M., 1997. A two way view of gender bias in medicine. J. Epidemiol. Community Health 51 (2), 106.

Santiesteban, P., Endres, M., Weimer, W., 2022. An analysis of sex differences in computing teaching evaluations. In: Proceedings of the Third Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering. In: GE@ICSE '22, Association for Computing Machinery, New York, NY, USA, pp. 84–87. http://dx.doi.org/10.1145/3524501.3527604, URL: https://dl.acm.org/doi/10.1145/3524501.3527604.

Saramäki, J., Kivelä, M., Onnela, J.-P., Kaski, K., Kertész, J., 2007. Generalizations of the clustering coefficient to weighted complex networks. Phys. Rev. E 75, 027105. http://dx.doi.org/10.1103/PhysRevE.75.027105, URL: https://link.aps.org/doi/10.1103/PhysRevE.75.027105.

Schimanski, L.A., Alperin, J.P., 2018. The evaluation of scholarship in academic promotion and tenure processes: Past, present, and future. F1000Research 7.

Scholar, G., 2023. Google Scholar. URL: https://scholar.google.com/.

Soffer, S.N., Vazquez, A., 2005. Network clustering coefficient without degree-correlation biases. Phys. Rev. E 71 (5), 057101.

Szlavi, A., Versino, S., Zanardi, I., Bolesta, K., Jaccheri, L., 2023. Bridging values: The inclusion of young generations in computing. In: Antona, M., Stephanidis, C. (Eds.), Universal Access in Human-Computer Interaction. Springer Nature Switzerland, Cham, pp. 154–170.

Tahsin, N., Ahmed, N.S., Asad, M., Sakib, K., 2022. Can female underrepresentation in information technology be solved through an awareness-based approach? In: Proceedings of the Third Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering. In: GE@ICSE '22, Association for Computing Machinery, New York, NY, USA, pp. 1–5. http://dx.doi.org/10.1145/3524501.3527606, URL: https://dl.acm.org/doi/10.1145/3524501.3527606.

Tahsin, N., Hasan, M.A., Islam, R., Sakib, K., 2023. Cognitive distance and women in software engineering: An empirical study in the context of Bangladesh. In: 2023 IEEE/ACM 4th Workshop on Gender Equity, Diversity, and Inclusion in Software Engineering (GEICSE). pp. 1–8. http://dx.doi.org/10.1109/GEICSE59319.2023.00005, URL: https://ieeexplore.ieee.org/abstract/document/10213433.

Trinkenreich, B., Britto, R., Gerosa, M.A., Steinmacher, I., 2022. An empirical investigation on the challenges faced by women in the software industry: a case study. In: Proceedings of the 2022 ACM/IEEE 44th International Conference on Software Engineering: Software Engineering in Society. In: ICSE-SEIS '22, Association for Computing Machinery, New York, NY, USA, pp. 24–35. http://dx.doi.org/10.1145/3510458.3513018, URL: https://dl.acm.org/doi/10.1145/3510458.3513018.

Wang, Y., Zhang, X., Wang, W., 2023. Fundamentalists, integrationists, & transformationists: An empirical theory of men software engineers' orientations in gender inequalities. In: 2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS). (ISSN: 2832-7616) pp. 25–36. http://dx.doi.org/10.1109/ICSE-SEIS58686.2023.00009, URL: https://ieeexplore.ieee.org/abstract/document/10173887?casa_token=p29o3o734oAAAAAA:E2YGtbdDEjyRo07ltQZMVB520WG85DnT2xqlgNzIC7UJCuf_rYGVn2lax4vep6s5EeUnppJK.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A., 2012. Experimentation in software engineering. Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-642-29044-2, URL: http://link.springer.com/10.1007/978-3-642-29044-2.

Zhao, X., Young, R., 2023. Workplace discrimination in software engineering: Where we stand today. In: 2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS). (ISSN: 2832-7616) pp. 188–193. http://dx.doi.org/10.1109/ICSE-SEIS58686.

2023.00026, URL: https://ieeexplore.ieee.org/abstract/document/10173897?casa_token=i5yovMNoAL0AAAAA:ZEhs09kLwkjGV6BZVksW5WNdoBp0tX5w-HXOs8gw1OmtDL5NdTiso_p0O_CaNoXdML-xNl_z.

Andrea D'Angelo is a Ph.D. student in Computer Science at the University of L'Aquila, Italy, where he previously earned his Bachelor's and Master's degrees in Computer Science in 2019 and 2022, respectively. His research interests revolve around information retrieval, Gender Fairness, and Natural Language Processing. He is a member of cost action EUGAIN (https://eugain.eu/) and published several papers on the topic of gender fairness.

Giordano d'Aloisio is a Ph.D. student in Computer Science at the University of L'Aquila. He obtained his bachelor degree in Computer Science for Business Economy at the University of Chieti-Pescara in 2017 and his master degree in Computer Science at the University of L'Aquila in 2021. From 2019 to 2024, he is a member of the Territori Aperti project where he is responsible of the Data Integration activity. He is also a member of the Emeliot, Fair-EDU, and PinKamP projects. His research is mostly focused on Data Science and Software Engineering techniques for the quality assurance of Machine Learning systems, with a particular attention on Bias and Fairness of ML algorithms.

Francesca Marzi is a computer scientist at the ICT Department of the University of L'Aquila, Italy, where she previously earned her Master's degrees and Ph.D. in Computer Science. Her main research interests are Data Science, Operations Research, Integer Programming and Combinatorial Optimization.

Antinisca Di Marco is Associate Professor in Computer Science at University of L'Aquila. Her main research topics are Software Quality Engineering, Data Science, Quality (such as fairness) in Learning Systems and in Data Science approaches, and Bioinformatics. She is involved in several national and international projects on such topics. She is the responsible of the research infrastructure of Territori Aperti project, co-PI of the SoBigData.it project and the director of the INFOLIFE CINI Laboratory node in L'Aquila. Since 2018, she is involved in several actions and projects aiming at improving equal opportunities in STEM (Science, Technology, Engineering and Mathematics). In particular, she is member of the cost action EUGAIN (https://eugain.eu/), and co-ideator and co-coordinator of PinKamP (www.pinkamp.disim.univaq.it).

Giovanni Stilo is an associate professor and the head of the Data Science Degree program at the Department of Information Engineering, Computer Science, and Mathematics at the University of L'Aquila. He received his Ph.D. in Computer Science in 2013, and 2014 he was a visiting researcher at Yahoo! Labs in Barcelona. Between 2015 and 2018, he was a researcher in the Computer Science Department at La Sapienza University in Rome.

His research interests are related to machine learning, data mining, reinforcement learning, and artificial intelligence, with a special interest in (but not limited to) trustworthiness aspects such as Bias, Fairness, and Explainability.

He also organized several international workshops and tutorials held in conjunction with top-tier conferences (AAAI, ICDM, CIKM, and ECIR). He is involved as an editor, reviewer, and PC of top-tier journals and conferences, such as TITS, TKDE, DMKD, AI, KAIS, AIIM, ICDM, KDD, CIKM, AAAI, and IJCAI.