Scientometrics (2021) 126:3807–3840
https://doi.org/10.1007/s11192-021-03885-3

Assessment of gender divide in scientific communities

Antonio De Nicola[1] · **Gregorio D'Agostino**[1]

Received: 24 February 2020 / Published online: 6 March 2021
© The Author(s) 2021

Abstract
Increasing evidence of women's under-representation in some scientific disciplines is prompting researchers to expand our understanding of this social phenomenon. Moreover, any countermeasures proposed to eliminate this under-representation should be tailored to the actual reasons for this different participation. Here, we take a multi-dimensional approach to assessing gender differences in science by representing scientific communities as social networks, and using data analytics, complexity science methods, and semantic methods to measure gender differences in the context, the attitude and the success of scientists. We apply this approach to four scientific communities in the two fields of computer science and information systems using the network of authors at four different conferences. For each discipline, one conference is based in Italy and attracts mostly Italians, while one conference is international in both location and participants. The present paper provides evidence against common narratives that women's under-representation is due to women's limited skills and/or less social centrality.

Keywords Gender · Social networks · Semantics · Complex networks

Introduction

Low participation of women in some of the STEM (Science, Technology, Engineering, and Mathematics) disciplines, such as computer science (CS) and information systems (IS), is a social phenomenon that is gaining the attention of the scientific community (EIGE 2019; Fatourou et al. 2019; Kersey et al. 2018; Leslie et al. 2015; Payton and Berki 2019; Reilly et al. 2019; UNESCO 2017). There are several reasons why public institutions such as the European Commission want to increase the participation of women in STEM disciplines: first of all in order to achieve gender equality and combat gender discrimination (Fatourou et al. 2019); secondly in order to meet a growing global demand for STEM professionals (especially in the digital sector); and finally, as revealed by a study by the European

✉ Antonio De Nicola
antonio.denicola@enea.it

Gregorio D'Agostino
gregorio.dagostino@enea.it

[1] Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Casaccia Research Centre, Via Anguillarese 301, 00123 Rome, Italy

Institute of Gender Equality (EIGE), because reducing the gender gap in STEM education would lead to more jobs and an increase in gross domestic product (EIGE 2019).

Although there are some attempts to explain this phenomenon, the reasons for male over-representation are still unclear and controversial (O'Dea et al. 2018). Currently, most of the studies on gender diversity mainly address differences of gender ratios of scientists writing papers or being the first authors (Mattauch et al. 2020; Holman et al. 2018; Larivière et al. 2013), who graduate in these fields, who have speaking opportunities at scientific events (Ford et al. 2018), who have been awarded the Nobel Prize (Lunnemann et al. 2019) or who hold relevant positions (e.g., Full Professor or manager of an IT company). For instance, the 2017 Taulbee survey (Zweben and Bizot 2018) of the Computing Research Association focuses on students of academic units in the United States and Canada in the information, computer science and computer engineering disciplines. According to it, women's bachelor's degrees in 2016/2017 were 19.2%, master's degrees were 29.6% and PhDs were 19.3%. However, the contribution of these reports based on gender ratios is limited as they do not take into account the nuances of reality, which embeds the specificity of the scientific field, the social relationships between people and the individual psychological features of the community members. Furthermore, (Huang et al. 2020) found that the increase in the rate of women over the last 60 years has gone hand in hand with an increase of gender divide in terms of productivity and impact.

The recent public debate on the reasons for the under-representation of women in STEM has suggested that they are due to intrinsic differences in aptitude rather than socio-cultural differences (Damore 2017). However, a study of Kersey et al. (2018) on mathematical skills of children revealed that males and females are equally equipped for mathematics during childhood. Then, a study of Reilly et al. (2019) showed small-to-medium gender differences for most individual countries but not globally. Moreover, (Leslie et al. 2015 and Payton and Berki 2019) presented the results of studies that provide evidence that under-representation of women in STEM is due to negative stereotypes claiming that success depends on raw and innate talent and that women do not possess it. All the aspects mentioned above reinforce the idea that gender differences in representation are mainly due to socio-cultural motivations rather than intrinsic differences.

A more in depth-analysis of gender diversity that goes beyond differences in gender ratios is needed. To this end, the scientific community is very active in presenting studies to assess the gender gap. By harnessing the power of available data on research publications, new insights on gender differences can be produced. For instance, (Duch et al. 2012) provided evidence that gender disparities in publication rate and impact are discipline-specific. According to the Elsevier report (Elsevier 2017) focusing on gender in the global research landscape, women are slightly more involved in interdisciplinary research and are generally less internationally mobile than men. Bendels et al. (2018) analyzed the proportion of female authorships as well as the proportion of female first and last authorships to get insights on gender disparities in high-quality research. The conclusion of their research shows that, compared to men, women are under-represented as authors of prestigious papers, especially in those attracting the highest citation rates. Based on a study on more than one million of computer scientists, (Jadidi et al. 2018) found that the success of a computer scientist is mainly due to specific collaboration patterns (e.g., tendency in collaborating with more colleagues, longer-lasting collaborations) and that women are less likely to adopt them.

Another approach to address gender diversity is to propose indices or frameworks to assess the gender gap. For instance, the gender diversity index proposed by the European project GEnder Diversity Impact - Improving research and innovation through gender

diversity (GEDII) aims at measuring gender diversity in research teams (Humbert and Guenther 2018) by focusing on the representation of women and men in the category that is either most desirable (e.g. a senior role in the team) or inclusive (e.g. people with care responsibilities) and the extent to which women and men can equally progress within functional levels. This index mainly relies on demographic indicators, such as age, marital status and education, and functional indicators, such as team role and contract. However, this approach requires surveying individual teams, does not address cognitive aspects such as influence, creativity, and authority and is not based on an accurate analysis of the social context and the discipline. Among the existing periodic reports, the global gender gap report by the World Economic Forum is published annually to capture the extent of gender-based inequalities and to track their progress over time. It proposes an index that examines the gender gap across four subindices: economic participation and opportunity, educational attainment, health and survival, and political empowerment. The global gender gap report of 2018 (WEF 2018) identifies Artificial Intelligence (AI) as a driver for innovation in the near future and presents an analysis of AI skills of LinkedIn users based on genders and geography. However, according to it, only 22% of the global professionals in AI are females. This undermines females' ability to be key players in developing AI methods and technology. Recently, the EU Commission published another report to assess gender disparities in the society (EC: Directorate-General for Justice and Consumers 2018). This examined female labour market participation and gender pay, earnings and pension gaps in Europe.

In this paper, we define gender divide as any gender difference in the careers of scientists, whether female scientists are disadvantaged or advantaged. The research questions in this paper concern gender divide and are the following: (RQ_1) Can we assess gender divide in a scientific community? As a group, are women and men (RQ_2) equally successful, (RQ_3) equally creative, (RQ_4) equally key to the larger community, (RQ_5) equally important in determining the future direction of research? As a group, do women and men (RQ_6) study different topics, (RQ_7) work with each other, (RQ_8) have different sources of inspiration? (RQ_9) Does the gender divide depend on discipline? (RQ_{10}) Within a discipline, does the gender divide vary on a geographical basis?

These research questions are addressed by proposing a methodological framework for gender divide assessment. This science of science approach (Fortunato et al. 2018) aims to be a multi-perspective tool to assess gender divide in research communities based on real data and leveraging techniques from complex network analysis (Newman 2010) and semantic analysis (Mika 2007). The present paper shares with (Tommaso et al. 2017; Di Tommaso et al. 2020) an analysis of gender diversity by means of social network analysis, but proposes different metrics. We apply these metrics to case studies of the conferences of four different scientific communities. Two of the communities research information systems, while the other two research computer science. For each discipline, both an Italian community and an international one were analyzed. The Italian conferences are ITAIS[1] and SEBD[2], while the international ones are ICIS[3] and VLDB[4]. The difference between ITAIS and ICIS, as well as that between SEBD and VLDB, is related to the attendants: there are

[1] Web site of the conference of the Italian Chapter of AIS (Association for Information Systems): http://www.itais.org.

[2] Web site of the Italian Symposium on Advanced Database Systems: http://www.sebd.org.

[3] Web site of the international Conference for Information systems: https://aisnet.org/page/ICISPage.

[4] Web site of the International Conference on Very Large Data Bases: http://www.vldb.org.

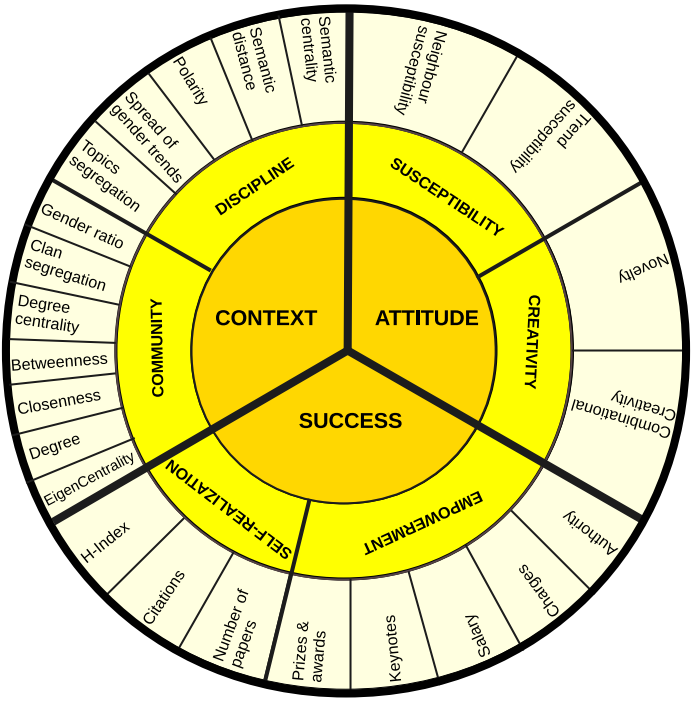

Fig. 1 Framework for gender divide assessment

mainly Italian scientists participating to ITAIS and SEBD, whereas scientists from all over the world attend ICIS and VLDB. While ITAIS and ICIS deal with topics related to the information systems discipline, SEBD and VLDB are concerned with computer science. Finally, it is noteworthy that the participants in these conferences are mainly professors and researchers. However, the ITAIS and ICIS participants mainly come from business schools and engineering universities in contrast to the SEBD and VLDB participants exclusively coming from engineering universities. Preliminary results concerning only the ITAIS community and an early version of the methodological framework for gender divide assessment were discussed in D'Agostino and De Nicola (2019). That study revealed that, despite the observed diversity, females play a significant role in this community. The results of the statistical analysis underlying this paper are presented in the Supplementary Information.

Methodological framework for gender divide assessment

Our methodological framework (Fig. 1) addresses gender divide along three dimensions. Each dimension is associated with a set of metrics and corresponding indices, which are described in Table 1. Data availability is likely to limit which of these indices can be measured.

The first dimension is the *context*, the environment within which the members of the community operate. The corresponding metrics are the *discipline* (Nielsen et al. 2018), describing the domain of interest under analysis, and the *community*, describing the

Table 1 Summary of the indices

Index	Description
Authority	Authority measures to what extent the topics in an author's papers influence the topics in his/her coauthors' papers.
Betweennes	Betweenness (B) (Freeman 1978) measures how important were a node if all of them would try to communicate along the network by the shortest path. That is, supposing anyone sends a message to anyone, how many of such messages pass through a node.
Charges	Number of positions of responsibility for controlling or caring for something. Examples could be the role of a scientist inside an organization (e.g., Full Professor) or a community.
Citations	Bibliographic index that measures the overall number of times a researcher has been cited by other scientists.
Clan segregation	Clan segregation indicates the presence in a community of prominent clusters of researchers (i.e. clans) consisting of either only males or only females.
Closeness	Closeness measures the average harmonic distance for a member to reach any other member of the community (Bavelas 1950).
Combinational creativity	Combinational creativity measures the ability of a researcher to combine different topics (Gero 2000). It is given by the number of times he/she has been the one that combined two topics for the first time.
Degree	Number of coauthors of a member the community.
Degree centrality	Degree centrality (D_{cen}) for a node is the fraction of nodes it is connected to. It is normalized by dividing by the maximum possible degree in a simple graph (n-1) where n is the number of nodes.
EigenCentrality	EigenCentrality (Bonacich 1987) can be interpreted as the probability of news to reach a node upon spreading on the network.
Gender ratio	Gender ratio allows to measure the female to male ratio.
H-index	H-index is defined as the maximum value of h such that the given researcher has published h papers that have each been cited at least h times (Hirsch 2005).
Keynotes	Number of *keynotes* held by a researcher.
Neighbour susceptibility	This index measures to what extent the topics in an author's papers are influenced by the topics in his/her coauthors' papers.
Novelty	Novelty measures the ability of a researcher to produce work that is both novel (i.e., original, unexpected) and appropriate (i.e., useful, adaptive concerning task constraints)" (Sternberg and Lubart 1999). It is given by the number of times he/she has been the one that introduced a topic for the first time.
Numbers of papers	Number of papers written by a researcher.
Polarity	Polarity captures the difference of attention paid by females and males to the topics. Polarity of a topic c_k is computed as following: $P(c_k) = P_m(c_k) - P_f(c_k)$ where $P_m(c_k)$ and $P_f(c_k)$ are the average semantic profiles concerning the topic c_k, respectively, of males and of females. Negative values of polarity correspond to the topics used mainly by females whereas positive values to those used mainly by males.
Prizes & awards	Number of prizes and awards won by a member of a community.
Salary	The salary received by a member of a community.
Semantic centrality	Semantic centrality measures to what extent the topics in an author's papers concern the key topics (i.e. the most popular ones) in the conference papers as a whole.
Semantic distance of genders	Semantic distance of genders allows to assess the diversity of semantic profiles relative to females and males.

Table 1 (continued)

Index	Description
Spread of gender trends	Spread of gender trends can be used to assess how focused the interests of each gender are.
Topics segregation	Topics segregation measures whether there are clusters of topics covered only by females or only by males.
Trend susceptibility	Trend susceptibility measures to what extent the topics in an author's papers are influenced by the topics in the conference papers as a whole.

features that emerge from existing social relationships and collaborations (Zeng et al. 2016) and the female to male ratio. The indices that were considered for the discipline metrics are based on semantic analysis. They are: *topics segregation* (Zeng et al. 2016), *spread of gender trends*, *polarity*, *semantic distance* and *semantic centrality*. The basic idea is that semantics, as well as natural language processing, can benefit quantitative social science (Garg et al. 2018). Topics segregation measures whether there are clusters of topics that are only covered by females or only by males. Spread of gender trends can be used to assess how focused the interests of each gender are. Polarity captures the difference in the attention paid by females and males to the topics. A semantic profile is the set of interests of a member of the social network together with the corresponding weights. A weight represents both the degree of interest in a topic and the estimation of the $L_i(c_k, t)$ likelihood of the h_i member to be interested in a c_k topic in year t (D'Agostino et al. 2015), which is computed as the relative frequency of the publications (these can be considered as "expressions of interest") of h_i in c_k. Thus, $L_i(c_k, t) = \frac{\nu_{h_i}(c_k,t)}{\sum_{c_k} \nu_{h_i}(c_k,t)}$, where $\nu_{h_i}(c_k, t)$ represents how many papers, written by the member h_i before the considered year t, are indexed by the c_k topic. This function, by definition, spans the [0, 1] range; the unitary value represents a total interest in the subject, while a null value means no interest at all. Semantic distance of genders allows to assess the diversity of semantic profiles related to females and males. Semantic centrality measures to what extent the topics characterizing an author's profile belong to the main-stream or to some minor subjects. This index therefore measures how key the topics treated by her or him are and to what extent the topics in an author's papers concern the key topics (i.e. the most popular ones) in the conference papers as a whole. The semantic analysis proposed for this metrics could be used to study the benefits of a greater involvement of females for scientific progress as done by Nielsen et al. (2017). Then, seven indices were considered for the community metrics. *Gender ratio* allows to measure the ratio between females and males whereas *clan segregation*, *degree centrality*, *betweenness*, *closeness*, *degree* and *eigencentrality* allow to measure the topological properties (Boccaletti et al. 2006) of the social network relating to the social inclusion of its members. Note that each scientific community was considered as a social network and, hence, as a graph where the nodes represent the authors and the edges the co-authorship relationships. Yang et al. (2019) presented evidence that network centrality indices are relevant for gender divide assessment as they can also be used to predict the success of an individual.

The second dimension of the framework is the *attitude*. This concerns the psychological tendencies of the members of the community. One of its possible metrics is *susceptibility*, distinguished in *neighbour susceptibility* and *trend susceptibility* indices (D'Agostino et al. 2015). The former measures to what extent the topics in an author's papers are influenced

Table 2 Summary of the main characteristics of the datasets for each scientific community: community (Com.), discipline (Disc.), number of papers (Pap.), period, area, number of males (Mal.), number of females (Fem.), number of undetermined (Undet.), total number of authors (Total), and ratio

Com.	Disc.	Pap.	Period	Area	Mal.	Fem.	Undet.	Total	Ratio
ITAIS	IS	852	2007–2016	Italy	709	418	–	1127	1:1.70
SEBD	CS	495	2007–2016	Italy	578	172	–	750	1:3.36
ICIS	IS	4348	1980–2016	World	1163	483	4056	5702	1:2.41
VLDB	CS	2590	1980–2016	World	844	160	2928	3932	1:5.28

by the topics in his/her coauthors' papers, while the latter to what extent the same topics are influenced by those in the conference papers as a whole. Another possible metrics is *creativity*, defined as "the ability to produce work that is both novel (i.e., original, unexpected) and appropriate (i.e., useful, adaptive concerning task constraints)" (Sternberg and Lubart 1999). This can be assessed by the *novelty* and the *combinational creativity* indices (Gero 2000). Even if assessing gender differences in creativity is controversial (Abraham 2016), these two indices can shed new light on the analysis of creativity of the members of a community.

The last dimension of the framework is the *success*, which is about achieving objectives. This is associated with two metrics: *self-realization* and *empowerment* to assess, respectively, fulfillment by oneself of individual possibilities and how much power and authority is granted to an individual. The empowerment metrics can be assessed, for instance, by means of the number of *prizes and awards* won, the number of *keynotes* held by a scientist, the *salary*, the *charges*, and her or his *authority* (D'Agostino et al. 2015). Authority measures to what extent the topics in an author's papers influence the topics in his/her coauthors' papers. Examples of charges could be the role of a scientist inside an organization (e.g., Full Professor) or a community. For instance, Lerback and Hanson (2017) presented evidence that females are less invited as referees in peer review for scholarly publications than males. Finally, the self-realization metrics can be assessed by means of bibliographic indices such as *H-index*, *citations*, and *number of papers*.

A summary of all indices is presented in Table 1.

Data

For each community, we collected data about publications including the titles, the publication years, and the authors from the web sites of the above-mentioned conferences, from the DBLP computer science bibliography[5], which provides open bibliographic information on major computer science journals and proceedings, and from SCOPUS. Data collection was done based on papers availability and with the goal of having datasets with the same time-period, at least at geographical level. We analyzed data by means of the methods and tools presented in the next Section. Table 2 presents a summary of the characteristics of the datasets. The authors were classified according to our ability to infer their gender from the

[5] DBLP computer science bibliography: https://dblp.uni-trier.de.

first name. If no gender could be inferred from the first name of the author, it was considered *undetermined* (see the column labelled by Undet. in Table 2).

Data is available at https://tinyurl.com/genderdivide-dataset.

Methods and tools

The novel method of analysis introduced in this paper regards a community of people as a semantic social network (SSN). This consists of an ontology representing the semantics of the domain of interest, a social network, and the actual interests of the community of members with their weights (D'Agostino et al. 2015; Mika 2007). An ontology is a conceptual model of a fragment of an observed reality that consists of interlinking concepts concerning a given domain of interest (De Nicola and Missikoff 2016).

The main steps of the method are the following. First (1) a domain ontology (De Nicola and Missikoff 2016) is extracted from a repository of raw data to represent the research topics of the discipline. The second step (2) is to identify the members by their names and determining the gender of each participant. Then, (3) some expressions of interests of the members in topics are inferred from their publications. To the purpose of these three steps, we analyzed the titles and the names of the authors of the papers presented at the conferences. We used lists of masculine and feminine names publicly available on the web[6] and we resorted to some pictures available on the web when the names were ambiguous. The third step allows to associate a dynamic (for each year) semantic profile to each member of the social network. The fourth step (4) is to make a topological analysis of the social network by means of complexity science methods and techniques. The fifth step (5) is to estimate susceptibility to trends and to neighbours and authority (Aral and Walker 2012; D'Agostino et al. 2015). This can only be done for a subset of community members that are here named semantically treatable. In fact, among all the authors, only a part of them published in at least two different years. Only those authors who exhibit a change of their interests during time are eligible for the analysis. For this purpose, we used a software application (namely the "attitude manager") developed on top of the interest propagation model and related equations of dynamics presented by D'Agostino et al. (2015). This model assumes that, as a person, each member tends to maintain her or his own beliefs; is partly influenced by others interacting with them (one-to-one interaction); and is partly influenced by trends (one to all interaction). The sixth step (6) is to identify creative members of the social network by detecting those that introduced new topics or novel combinations of topics. The last step is to measure success indices depending on data availability (e.g., number of papers, citations, ...).

The above-presented method is supported by a tools suite that we developed for the purpose. This takes as input natural language texts from the conference papers and performs the analysis of the semantic social network. Figure 2 shows a sketchy representation of the architecture of the tool suite that consists of five modules that are presented as follows.

The programming languages adopted to build the software are Java and Python. We wrote the code of most of the modules from scratch; however, we used also several existing libraries. Among them we cite: Apache Lucene, Colt, CommonsMath, rdf4j, for JAVA, and NetworkX and matplotlib, for Python.

[6] See, for instance, https://www.ssa.gov/oact/babynames/decades/century.html.

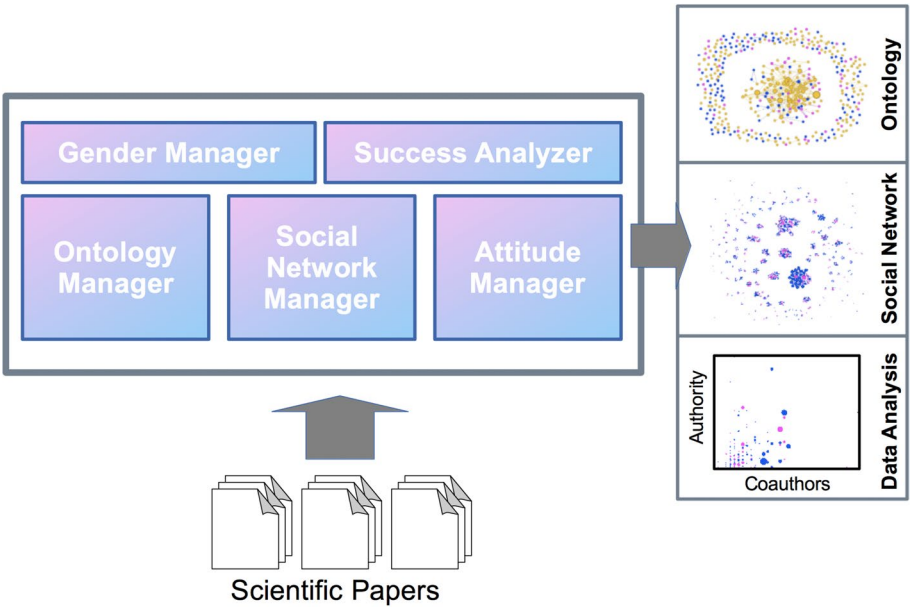

Fig. 2 Tool suite architecture for gender assessment of scientific communities

The ontology manager takes as input the available titles of the papers. We limited our analysis to the titles as it is possible to extract from them the main topics treated in the papers. There are other possible choices for the semantic analysis such as the use of the keywords. However, it is worth noting that sometimes authors' keywords do not refer to the specific content of the paper, but to the general state of the art in the field and, hence, they could include topics not related to the main subject of the paper. The ontology manager automatically extracts a set of topics and, by following the incremental ontology engineering methodology UPON lite (De Nicola and Missikoff 2016), a set of specialization relationships and a set of generic relationships between them. This module implements the following workflow to produce an approximation of the set of basic topics.

1. Replacing HTML tags with characters (e.g., "Ë" with "Ë").
2. Removing titles containing words that clearly indicate that the titles do not refer to scientific papers (e.g., editor in chief, poster).
3. Removing not-English titles by means of the language-detection techniques available with Apache Lucene[7] Java library.
4. Tokenizing titles to split titles in tokens by managing punctuation and function words (e.g., is, can, then).
5. Stemming to reduce words to their lexemes. A lexeme is a unit of lexical meaning that exists regardless of the number of inflectional endings it may have or the number of words it may contain (Hughes 2004). Accordingly, terms such as "innovation" and "innovations", or "topological" and "topology", correspond to the same lexeme.
6. Computing frequencies of lexemes and multi-words lexemes.

[7] Apache Lucene: https://lucene.apache.org/core.

7. Identifying the minimum annotation set of topics. These correspond to the single and multi word lexemes that permit to index all the papers of each dataset by means of at least one (single or multi word) lexeme.

We manually checked the identified topics to assess their quality. Precision of the ontology extraction process is 92%. Specialization relationships were identified by means of linguistic patterns (e.g., project manager IS_A manager) and generic relationships by means of co-occurrence of terms. The ontology was automatically saved in RDF (Resource Description Framework) and GML (Graph Modeling Language) for visualization purposes. Then, together with the gender manager, this module allowed to perform the gender-based assessment of the semantics of the field.

The gender manager module takes as input a list of masculine and feminine given names and determines the gender of the author. Given names of authors that do not belong to the above-mentioned lists should be manually checked to identify the gender.

The social network manager extracts the social network by analyzing the authors of the publications and builds a graph. If two scientists are co-authors of at least one paper, the social network manager builds a connection in the graph. With regard to the homonymy issue, data about SEBD, ICIS, and VLDB communities were collected from the DBLP computer science bibliography. An existing study (Kim 2018) shows that DBLP assigns author names quite accurately to distinct authors, possibly due to its hybrid disambiguation approach combining algorithmic disambiguation and manual error correction. Concerning ITAIS, we collected data from different sources (i.e., SCOPUS, conference agendas and conference proceedings) because not all data were available from DBLP. However, as the ITAIS community is not very large, we were able to check all the persons one by one and we did not detect any cases of homonymy. Then, the social network manager computes some topological features of the network, such as closeness, betweenness, eigencentrality, and degree.

The attitude manager estimates the susceptibility indices of treatable members of the social network by means of an algorithm based on information diffusion (D'Agostino et al. 2015). Then it detects the community members who introduced novel topics or novel combinations of them for each year of the time interval considered.

The success analyzer performs the analysis of the success. In particular, it estimates authority values for treatable members of the social network. Authority of a member is inferred from the susceptibility values of connected treatable members.

Context

As already mentioned, the first dimension is the context. This is associated to the discipline and the community metrics.

Discipline

To the purpose of discipline assessment, we considered three indices for the ITAIS and SEBD scientific communities: topics segregation, polarity, and semantic centrality. Instead, we considered only semantic centrality for the ICIS and VLDB scientific communities, due to our limited ability in recognizing gender from international first names.

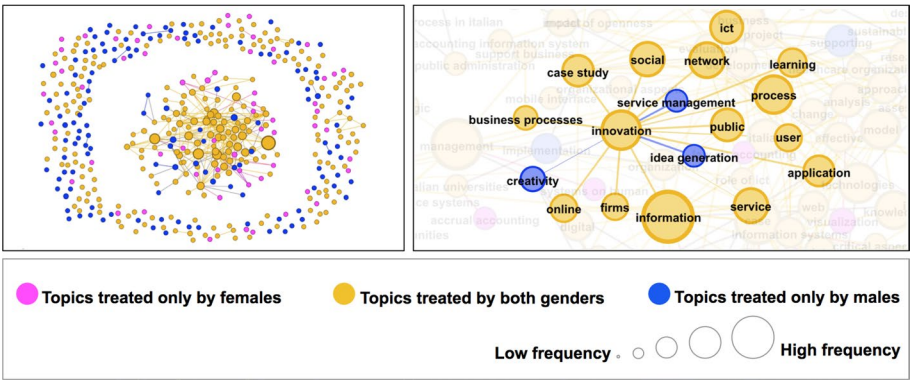

Fig. 3 Information systems ontology and an excerpt from its most connected part as resulting from the ITAIS dataset

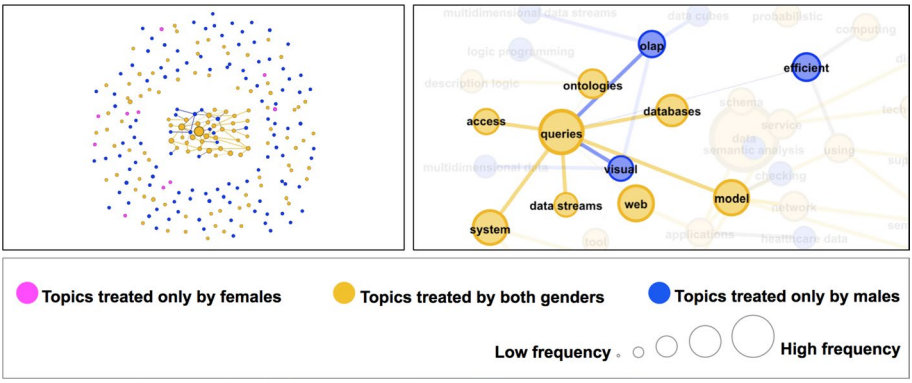

Fig. 4 Computer science ontology and an excerpt from its most connected part as resulting from the SEBD dataset

Topics segregation measures whether there are topics treated exclusively by females or others only by males. This is an accurate means for early detection of ghettoisation (Nielsen et al. 2018) in disciplines. We identified 431 topics related to information systems for ITAIS. Figure 3 depicts the corresponding ontology concepts as a semantic network (Sowa 2006). These are linked together by generic relationships automatically detected by means of natural language processing techniques. The assumption here is that two topics are related if they co-appear in the whole set of titles more than twice. The left part shows the overall ontology, whereas the right part is limited to an excerpt from the most connected part concerning the topic *innovation* and the other topics directly connected with it. There are 127 topics exclusive of males, 62 treated by females only and 242 shared ones. While there are gender-specific topics, we did not observe clusters of topics treated only by males or only by females. This means that it is not possible to determine a subset of the IS domain specific for males or females. Similarly, the analysis of the SEBD community did not reveal the topics segregation issue (see Fig. 4).

Polarity allows to assess the gender-based polarity of the topics and, hence, if a topic is used more by females than males or viceversa. Polarity of a topic c_k is computed as

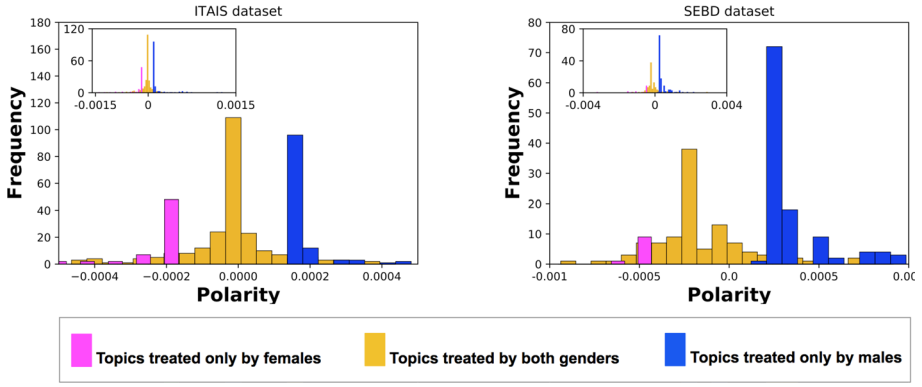

Fig. 5 Polarity distribution of topics. (Color figure online)

following: $P(c_k) = P_m(c_k) - P_f(c_k)$ where $P_m(c_k)$ and $P_f(c_k)$ are the average semantic profiles concerning the topic c_k, respectively, of males and of females. Negative values correspond to the topics used mainly by females, whereas positive values to those used mainly by males. The left part of Fig. 5 shows the gender-based polarity of the topics for the ITAIS community. Yellow rectangles correspond to topics used by both genders; blue to those used only by males; and purple to those used only by females. Most topics have polarity close to 0. Thus, they are indiscriminately used by both males and females. Conversely, in the case of the SEBD community, there is a clear evidence from the right part of Fig. 5 that there are several topics with polarity higher than 0. This means that these concepts are used more by males than by females. Hence, this analysis reveals that there is a polarity issue in the SEBD community but not in the ITAIS one.

Semantic centrality measures to what extent the topics in an author's papers belong to the main-stream or to some minor subjects. The semantic centrality of the h_i author is defined as following:

$$\mathcal{SC}_{h_i} = \sum_{c_k} L_{h_i}(c_k) \cdot L_s(c_k); \tag{1}$$

where $L_{h_i}(c_k)$ is her or his (their hereafter) degree of interest in the c_k topic and $L_s(c_k)$ is the likelihood for the environment to provide some information on the c_k topic (D'Agostino et al. 2015). This index spans the [0,1] range. According to the analyzed data (see the Supplementary Information), ITAIS is the only community where the female role is more central. In fact, $\Delta_{\bar{\mathcal{SC}}_{ITAIS}} = \bar{\mathcal{SC}}_M$-$\bar{\mathcal{SC}}_F = -0.00015$. Conversely, the role of males is more central in the SEBD ($\Delta_{\bar{\mathcal{SC}}_{SEBD}} = +0.0004$), ICIS ($\Delta_{\bar{\mathcal{SC}}_{ICIS}} = +0.0004$), and VLDB ($\Delta_{\bar{\mathcal{SC}}_{VLDB}} = +0.00019$) communities.

Community

The first index that we considered to the purpose of community assessment is gender ratio. The ITAIS and SEBD scientific communities consist, respectively, of 1127 authors and 750 authors. In both cases, female representation is less than the average as only 37.09% of authors are females in ITAIS (418 people) and 22.93% (172 people) in SEBD. Hence, males are 62.91% (709 people) in ITAIS and 77.07% (578 people) in SEBD. This means

that, in line with the trend for science in Italy, the involvement of women in this sector is certainly incomplete. The question is whether the role played by women is as significant as that of males. Concerning the ICIS and VLDB communities, we recognized the gender from the first names only of a limited number of authors. In particular, in ICIS, we recognized 483 females and 1163 males, while the gender was undetermined for 5702 authors. Among recognized authors, here, females are 29.34% and males 70.66%. Similarly, in VLDB, we recognized 160 females and 844 males, while the gender was undetermined for 3992 authors. Among the recognized authors, here, females are 15.94% and males 84.06%. For both the information systems and computer science disciplines, we observed that the percentage of female authors is higher in the Italian scientific communities (ITAIS and SEBD) than in the international ones (ICIS and VLDB). Moreover, we observed that the percentage of female authors is lower in computer science than in information systems.

As mentioned in the ''Methodological Framework for Gender Divide Assessment'' section, for the following community assessment indices, each scientific community was considered as a graph where the nodes represent the authors and the edges the co-authorship relationships. It is noteworthy that the identified scientists and the connections between them are not all those that exist. Taking into account other data would require a multiplex structure (D'Agostino and De Nicola 2016) where each layer represents the social network resulting, for instance, from a different conference or journal. Such analysis would probably provide better results but there are some issues that need to be further investigated (e.g., different conferences could concern slightly different disciplines). However, we kept them out-of-scope in this paper. Figure 6 provides a graph representation of the overall communities: authors are represented by balls, which sizes are proportional to the degree (i.e. number of coauthors) and colors depend on gender: females dress purple, while males dress blue, undetermined ones in ICIS and VLDB dress grey.

Another relevant feature of a community is its connectivity. For instance, 213 clusters are observed in ITAIS, which are here shortly referred to as "clans". A clan is a set of authors where each pair shares at least one publication. From the mathematical point of view, co-authorship represents an equivalence relation and clans are the resulting classes. It is possible to observe 7 prominent clans (with clan size $cs > 19$) in ITAIS, linking together several members, and several smaller ones, linking together smaller groups. The former represent well-connected members of the community; while the latter gather authors that wrote either few papers or with a restricted group of colleagues. Even if the gender ratio is lower than 1 (see Table 2), the above-mentioned prominent clans do not consist of only males or females (see Fig. 6 and Supplementary Information). ITAIS is the only community where the number of prominent clans is relevant. Table 3 presents some statistics about the identified clans in all the communities. In details, in the table, we included the number of researchers in the community (Persons), the number of clans (Clans), the number of clans with clan size (cs) higher than 19, the size of the largest clan (cs_{MAX}), the average clan size ($\bar{cs}$), the standard deviation (σ_{cs}), and the *normalized community entropy* $\hat{H}_c$. We define $\hat{H}_c$ as

$$\hat{H}_c = \frac{H_c}{H_{c,max}}, \tag{2}$$

where $H_c = -\sum_i (p_{cl_i} \cdot \ln p_{cl_i})$ is the *community entropy*; p_{cl_i} is the probability of a researcher to belong to the clan i; and $H_{c,max}$ is the maximum entropy of the community. The latter is the entropy of a completely disconnected community where every clan consists of only one researcher. The normalized community entropy spans the [0,1] range and

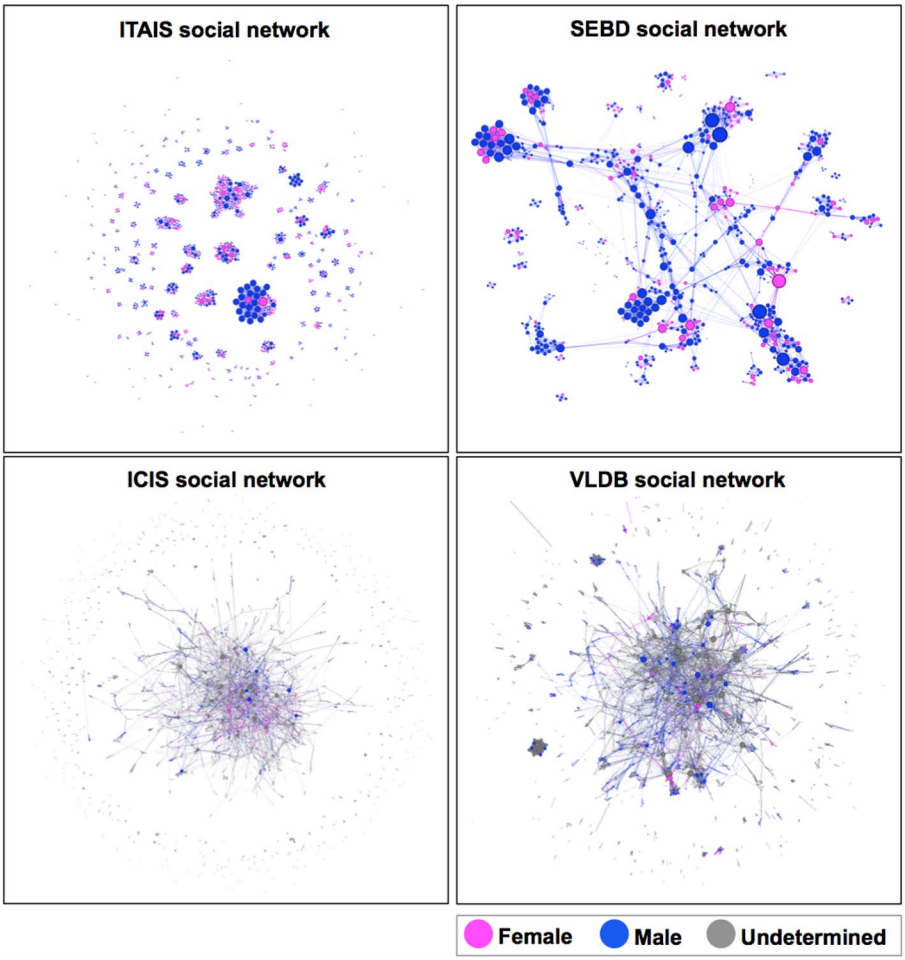

Fig. 6 Graph representation of the ITAIS, SEBD, ICIS, and VLDB social networks

Table 3 Summary of data concerning clans for each scientific community: community, overall number of persons (Persons), number of clans (Clans), number of clans with clan size (cs) higher than 19, maximum clan size, average clan size ($\bar{cs}$), standard deviation (σ_{cs}) of clan size, and normalized community entropy ($\hat{H}_c$)

Community	Persons	Clans	Clans ($cs > 19$)	cs_{MAX}	$\bar{cs}$	σ_{cs}	$\hat{H}_c$
ITAIS	1127	213	7	175 (15.53%)	5.29	13.29	0.64
SEBD	750	62	1	576 (76.80%)	12.10	72.25	0.21
ICIS	5702	699	3	3986 (69.90%)	8.16	150.58	0.29
VLDB	3932	407	4	2833 (72.05%)	9.66	140.15	0.26

Table 4 Comparison of the average values of the main centrality indices

Community	Gender	$\bar{B}$	$\bar{D}_{cen}$	$\bar{Cl}$	$\bar{D}_{all}$	$\bar{E}$
ITAIS	Females	0.000089	0.0029	0.0107	3.25	0.0012
	Males	0.000095	0.0035	0.0093	3.91	0.0061
SEBD	Females	0.0033	0.0081	0.102	6.10	0.0075
	Males	0.0037	0.0082	0.113	6.17	0.0061
ICIS	Females	0.00031	0.00062	0.085	3.56	0.0039
	Males	0.00044	0.00068	0.083	3.90	0.0042
VLDB	Females	0.00031	0.00114	0.098	4.48	0.000003
	Males	0.00063	0.00143	0.104	5.62	0.000916

measures the "state of disorder" of the community: 0 if all the researchers belong to only one clan and 1 if all the researchers belong to different clans. The number of clans with size higher than 19, $\bar{cs}$, and $\hat{H}_c$ in Table 3 reveal that ITAIS is the "less ordered" community where the largest clans have a similar size. Instead, SEBD, ICIS, and VLDB are characterized by a large clan and several other ones with negligible size. Furthermore, we analyzed the number of females in the largest clan of all the communities and we observed a significant clan segregation against females only in SEBD and a weakly significant one in VLDB. We did not observe any clan segregation in ICIS whereas, from a statistical perspective, females are better positioned than males in ITAIS with respect to this index.

Given the social network, several indices can be calculated to quantify the centrality of researchers in the considered scientific communities (Boccaletti et al. 2006). Table 4 reports some of those topological indices for the communities and the gender subsets.

Each of the reported indices carries information on the role played by females (or males) in a specific respect. Most of such indices make an appropriate sense for fully connected networks. Here, the general indices were evaluated by performing an average over all clans.

Betweenness (B) (Freeman 1978) measures how important were a node if all of them would try to communicate along the network by the shortest path. That is, supposing anyone sends a message to anyone, how many of such messages pass through a node. Even if males are more central than females in all the addressed scientific communities, this difference is not statistically significant in ITAIS and SEBD (see the Supplementary Information). Therefore, some gender divide in favor of males concerning betweenness was observed only in the international conferences.

Degree centrality (D_{cen}) for a node is the fraction of nodes it is connected to. It is normalized by dividing by the maximum possible degree in a simple graph (n-1) where n is the number of nodes[8]. Even if males are more central than females with respect to degree centrality in all the addressed scientific communities, this difference is not statistically significant in SEBD that, hence, is the only community where such type of gender divide in favor of males was not observed.

On the other side females exhibit an average higher closeness centrality (Cl) (Bavelas 1950) in ITAIS and ICIS. This means that the average harmonic distance for females to reach any other member of the community is longer than the same quantity for males.

[8] Degree centrality description was retrieved from NetworkX online reference documentation: https://networkx.github.io/documentation/networkx-1.9/reference/generated/networkx.algorithms.centrality.degree_centrality.html.

Closeness of males is higher in the other communities. Hence, such type of gender divide in favor of males was observed only in the computer science communities.

As can be seen at glance looking at the balls in Fig. 6, the average size for females is similar to the one for males in ITAIS. Actually, average degree (D_{all}) for females is 3.25, while for males it is 3.91. On average, females have about half collaborator less than males. This means that, within the ITAIS community, females tend to publish with about 20% fewer co-authors than males. Degree is higher for males in the other communities as well, but such gender divide in favor of males is not statistically significant in SEBD.

Eigencentrality (E) (Bonacich 1987) can be interpreted as the probability of news to reach a node upon spreading on the network. The results for the four case studies are a bit controversial as gender divide in favor of males was observed in the Italian community of information systems (ITAIS) and in the international one of computer science (VLDB) whereas it was not observed in the others (SEBD and ICIS).

Some community indicators are dependent as revealed by Martin-Hernandez (2013).

Attitude

The analysis of the attitude dimension addressed some individual characteristics of the members of the community, such as susceptibility to neighbours and trends, and creativity.

Susceptibility

Susceptibility indices are based on the idea that some psychological attitudes of members can be assessed by studying how the level of interest of authors in topics over time tracks the scientific discourse of the communities (Gerow et al. 2018). To this purpose, we considered a community as a semantic social network and we analyzed the diffusion of interests among the members by means of the approach presented by D'Agostino et al. (2015), which allows to estimate neighbour susceptibility (x_i) and trend susceptibility (x_{is}) for each member h_i.

Neighbour susceptibility measures the extent that topics in an author's papers are influenced by topics in coauthors' papers. Specifically, it measures how strongly correlated changes in an author's semantic profile are with changes in their co-authors semantic profiles. In contrast, trend susceptibility measures the impacts of the entire community. Specifically, it measures how correlated changes in an author's semantic profile are with the average changes in the semantic profiles of all conference papers. The formulas to calculate these indices are presented in the appendix.

As already mentioned in the "Data" section, these indices were estimated only for semantically treatable authors (h_i) that have written at least two papers in two different years. There are 347 (30.79%) semantically treatable authors in ITAIS, 280 (37.33%) in SEBD, 1771 (31.06%) in ICIS, and 1173 (29.83%) in VLDB. The general tendency for authors in all the communities is to be more influenced by trends than by coauthors. In fact, the average susceptibility to neighbours and to trends are, respectively, 4.1%, and 18.8% in ITAIS; 9.1%, and 14.1% in SEBD; and 9.8%, and 11.5% in ICIS. The average susceptibility to neighbours is higher than that to trends (7.1%, against 6.1%) only in VLDB. As can be seen in Table 5 and in the Supplementary Information, the difference between the average susceptibility to neighbours ($\bar{x}$) for males and that for females is significant only in SEBD where it is 7.1% for males and 16.2% for females. Hence,

Table 5 Comparison of the average values of the susceptibility indices and authority for semantically treatable values. Average susceptibility to neighbours: $\bar{x}$; average susceptibility to trends: $\bar{x}_s$; average authority: $\bar{a}$

Community	Gender	$\bar{x}$	$\bar{x}_s$	$\bar{a}$
ITAIS	Females	0.045	0.184	0.078
	Males	0.039	0.192	0.092
SEBD	Females	0.162	0.124	0.215
	Males	0.071	0.145	0.225
ICIS	Females	0.102	0.095	0.173
	Males	0.103	0.129	0.217
VLDB	Females	0.066	0.052	0.119
	Males	0.060	0.059	0.147

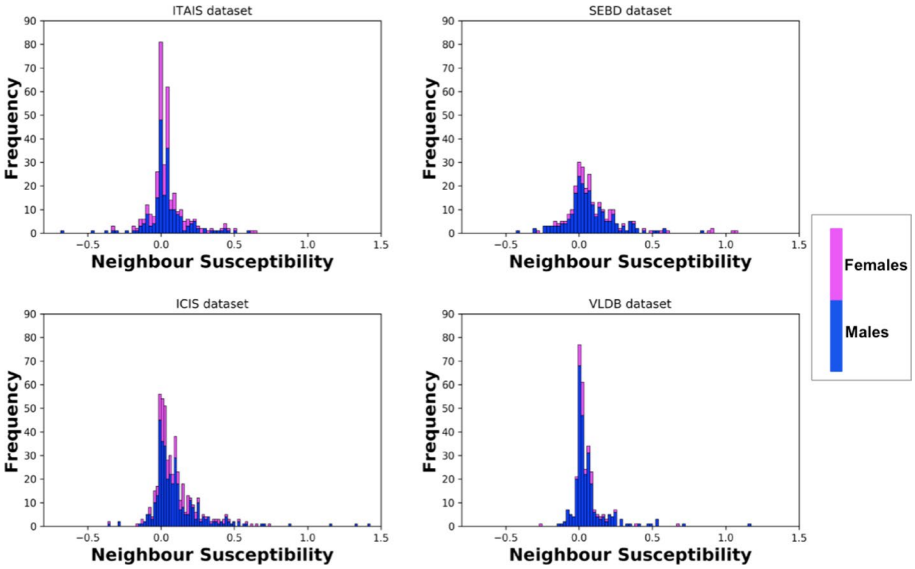

Fig. 7 Histograms of neighbour susceptibility

there is not gender divide concerning this index in the other communities. The difference between the average susceptibility to trends ($\bar{x}_s$) for males and that for females is not significant only in ITAIS. In the other communities, there is gender divide as the average susceptibility to trends for males is higher than that for females: 14.5% against 12.4% in SEBD, 12.9% against 9.5% in ICIS, and 5.9% against 5.2% in VLDB (see Table 5). Figures 7 and 8 depict, respectively, the histograms of neighbour and trend susceptibility concerning, respectively, the ITAIS, the SEBD, the ICIS, and the VLDB communities. Both the figures highlight the above mentioned differences of susceptibility indices between genders.

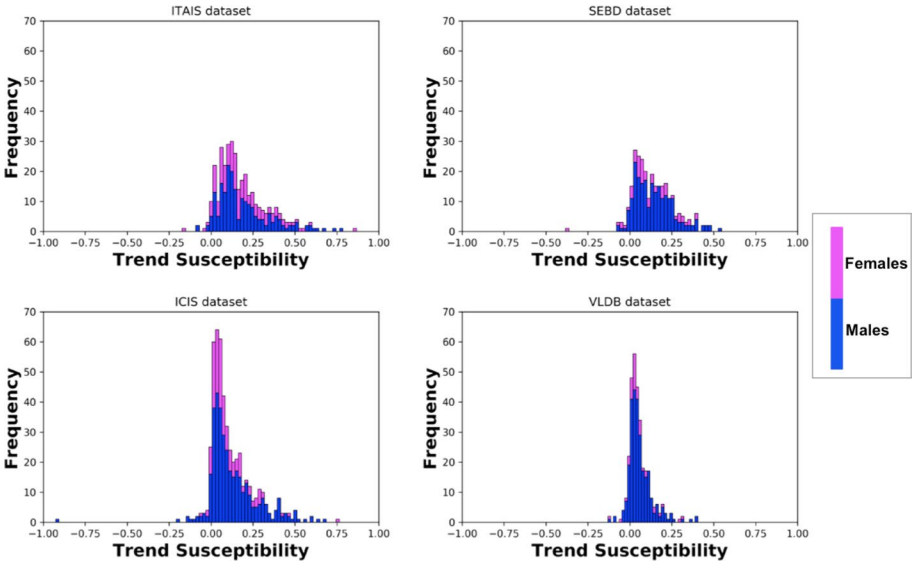

Fig. 8 Histograms of trend susceptibility

Creativity

Creativity in conceiving new ideas and technology is one of the elements required to disrupt science and technology. While Wu et al. (2019) assessed the role of the size of research teams in disrupting them, this study focuses on individual creativity by proposing two novel indices: novelty and combinational creativity. These permit to consider also the original work of scientists even if its importance is not recognized for several years after its publication, as happens with the so called sleeping beauties (Ke et al. 2015). As a preliminary step towards the definition of these two indices, for each year of the studied period, we identified the authors who first introduced a new topic in the community. Then, we considered combinational creativity, which is the ability to combine different topics (Gero 2000). Indeed, according to Uzzi et al. (2013), the highest impact in science is primarily due to conventional and unusual combinations of prior work. Along this line, we identified the authors that combined different existing topics first. Then, we compared the ratios of creative authors by gender with the general gender distribution of authors. Table 6 presents the results of this preliminary analysis. According to them, it was observed that there are more creative females in ITAIS and VLDB and less creative females in SEBD. In ICIS, females are more creative, from the novelty perspective, and less from the combinational creativity one. However, as such deviations are small and counting the number of creative authors does not take into account to what an extent an author is creative, we defined two new indices as it follows.

The novelty index $\mathcal{N}$ measures to what extent an author contributes to the formation of the discipline by introducing novel topics. The novelty index of the author h_i is defined as following:

Table 6 Results of the creativity analysis and average number of papers for the ITAIS, SEBD, ICIS, and VLDB communities. $\mathcal{N}$ authors are the authors introducing novel topics. $\mathcal{C}$ authors are the authors introducing novel combinations of topics. $\bar{\mathcal{N}}$ is the average novelty index. $\bar{\mathcal{C}}$ is the average combinational creativity index. $\bar{p}$ is the average number of papers

Community	Gender	Ratio (%)	$\mathcal{N}$ authors (%)	$\mathcal{C}$ authors (%)	$\bar{\mathcal{N}}$	$\bar{\mathcal{C}}$	$\bar{p}$
ITAIS	Female	37.09	37.88	37.38	0.00099	0.0010	2.04
	Male	62.91	62.12	62.62	0.00082	0.0008	1.90
SEBD	Female	22.93	21.78	22.22	0.00128	0.00123	2.27
	Male	77.07	78.22	77.78	0.00135	0.00136	2.34
ICIS	Female	29.34	31.20	29.19	0.000170	0.000153	1.89
	Male	70.66	68.80	70.81	0.000163	0.000162	2.07
VLDB	Female	15.94	16.17	16.46	0.00021	0.000208	1.68
	Male	84.06	83.83	83.54	0.00030	0.000282	2.16

$$\mathcal{N}_{h_i} = \frac{1}{N_{c_k}} \cdot \sum_{c_k} \frac{\delta_{k,i}}{n_{c_k}}, \tag{3}$$

where $\delta_{k,i} = 1$ if the author h_i was one of the first authors to propose the topic c_k; $\delta_{k,i} = 0$ if the author h_i was not one of the first authors to propose the topic c_k; n_{c_k} is the number of the authors that first proposed the topic c_k; and N_{c_k} is the overall number of topics. This index spans the [0,1] range.

Similarly, the combinational creativity index $\mathcal{C}$ measures to what extent an author contributes to the formation of the discipline by introducing novel combinations of topics. The combinational creativity index of the author h_i is defined as following:

$$\mathcal{C}_{h_i} = \frac{1}{N_{p_{kl}}} \cdot \sum_{p_{kl}} \frac{\delta_{p_{kl},i}}{n_{p_{kl}}}, \tag{4}$$

where p_{kl} is the pair of topics (c_k, c_l); $\delta_{p_{kl},i} = 1$ if the author h_i was one of the first authors to propose p_{kl}; $\delta_{p_{kl},i} = 0$ if the author h_i was not one of the first authors to propose p_{kl}; $n_{p_{kl}}$ is the number of the authors that first proposed p_{kl}; and $N_{p_{kl}}$ is the overall number of detected pairs of topics. Also this index spans the [0,1] range.

According to the $\bar{\mathcal{N}}$ and $\bar{\mathcal{C}}$ indices (see Table 6, Fig. 9, and the Supplementary Information for the detailed statistical analysis), females are more creative in the ITAIS community and less in the VLDB community from both the novelty and combinational creativity perspectives. Figure 9 depicts the scatter plots, concerning the ITAIS, the SEBD, the ICIS, and the VLDB communities, with the relationship between the novelty and the combinational creativity values. Several purple dots representing females are either on the right or on the top of the populated areas of the ITAIS scatter plot as well as they are almost absent in the same areas of the VLDB scatter plot. Moreover, it was observed some gender divide in favor of males from the combinational creativity perspective but not from the novelty perspective in the SEBD and ICIS communities. This is also revealed by observing that a limited number of purple dots are on the top of the populated areas of the SEBD and ICIS scatter plots of Fig. 9, while more purple dots are on the right of the populated areas.

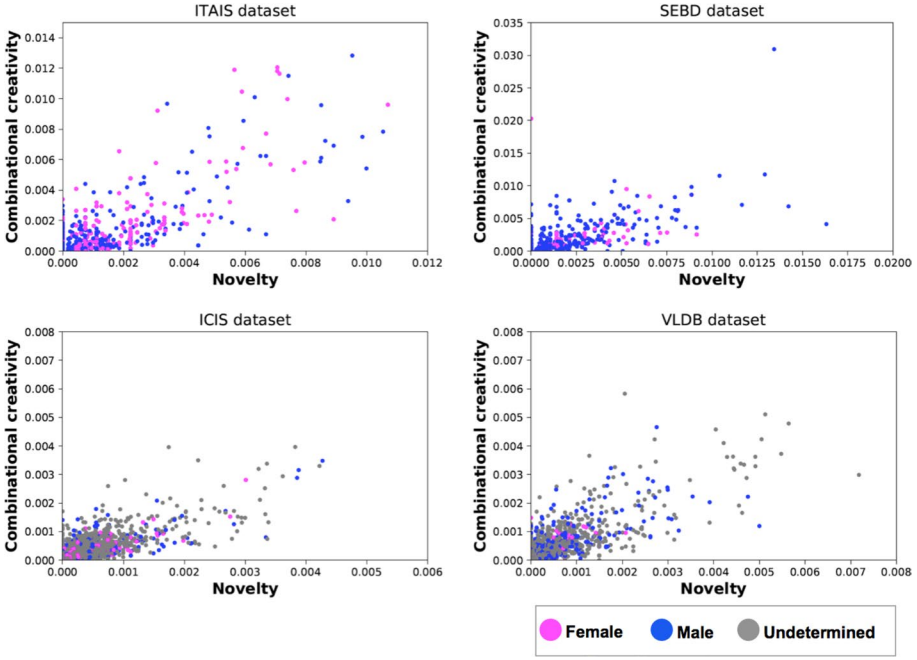

Fig. 9 Relationship between novelty and combinational creativity

Success

The analysis of the success dimension addressed the accomplishment of the scientist objectives.

Empowerment

Following the analytical approach presented by D'Agostino et al. (2015), we estimated the authority of semantically treatable authors (h_i) in the four communities by adding the neighbour susceptibility values (x_i) of their neighbours (i.e. coauthors) h_j: $a_i \stackrel{def}{=} \sum_{h_j \in N_{h_i}} x_j$, where N_{h_i} is the overall set of the coauthors of the author h_i. Table 5 presents the average values ($\bar{a}$) for females and males. The general tendency for authors in all the communities is that the average authority of males is higher than that of females. However, the observed differences of average values in the SEBD community are not significant from a statistical point of view (see Table 5 and the Supplementary Information). Figure 10 presents the histograms of authority distributions concerning the four communities. Again blue bins represent males, while purple ones represent females. This shows that the SEBD community is the only one where females are characterized by authority values that tend to be uniformly distributed in the whole range with respect to the other communities, where females are characterized mainly by low values. Figure 11 presents the relationship between the authority of authors and the number of published papers. Balls are proportional to the degree (i.e. number of coauthors). Again, blue balls represent males and purple ones represent females.

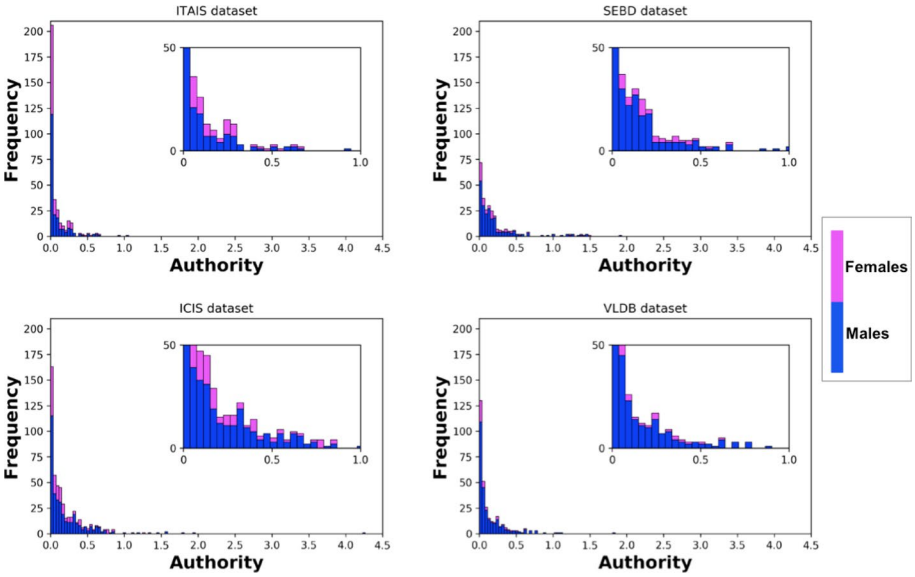

Fig. 10 Histograms of authority

The scatter plots in Fig. 11 reveal that, even if there is not a clear evidence of relationship between the authority and the number of papers as well as with the degree, the trend is that the degree affects the authority more than the number of papers. This can be noted, in particular, in the ICIS scatter plot. Hence, a suggestion to improve the authority values for females is to increase the number of their peers rather than their scientific production. Finally, Fig. 12 presents the scatter plots with the relationship between the trend susceptibility and the neighbour susceptibility for authors having authority greater than the average authority. Here, the size of the balls representing the authors is proportional to the authority. There is not a clear-cut evidence of any relationship between authority and one of the two susceptibility indices.

Self-realization

The last index that was considered for the analysis of the four communities is the average number of papers $\bar{p}$. ITAIS is the only community where the average value of this index for females is higher than that of males, while in the others males tend to publish more than females. All these results are statistically significant (see the Supplementary Information).

Statistical analysis for gender divide assessment

This section describes the statistical analysis that we performed to decide if there is gender divide related to an indicator. The statistical analysis may be regarded as testing the hypothesis of genuine gender divide against the "null hypothesis" that no gender divide is found. In some cases it provides evidence for a gender divide, while in other cases the set of females appears close similar to a random selected one.

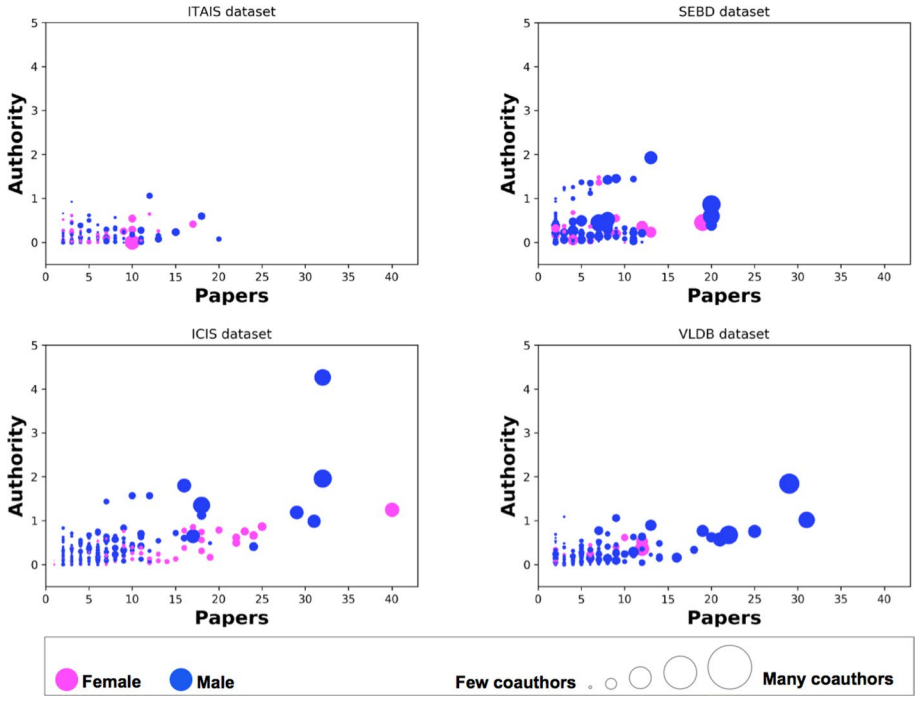

Fig. 11 Relationship between authority, number of published papers, and coauthors (ball size)

For each indicator, we present a figure showing the distributions of the average value of random samples of the same size of the female group (see Supplementary Information). Figure 13, for instance, presents the distributions of the sample averages of the combinational creativity index for the four scientific communities. This allows to estimate the *degree of belief* that the observed deviations for the female group are not accidental. The average values of combinational creativity index characterizing females and males are also outlined. We generated 100,000 samples by randomly selecting authors from the population of each available community. To this purpose, we implemented an algorithm in Python, which is presented in the following in pseudocode. The core of this algorithm is the `randrange` function of the Python module named `random`. `randrange` uses the Mersenne Twister algorithm (Matsumoto and Nishimura 1998) as core generator of pseudo-random numbers. In general, the vertical pink line, marked by an F, represents the average value of the indicator for females ($\bar{ind}_F$); the blue line, marked by an M, represents the average value for males ($\bar{ind}_M$); the grey line, marked by U, represents the average value for authors (in the ICIS and VLDB communities only) with undetermined gender ($\bar{ind}_U$); finally, the black line marked by R represents the average value for the above mentioned random samples ($\bar{ind}_R$). The former average values for the different groups are also reported in tables for all the considered communities (see, for instance, Table 7 for the combinational creativity and Supplementary Information for the other indicators).

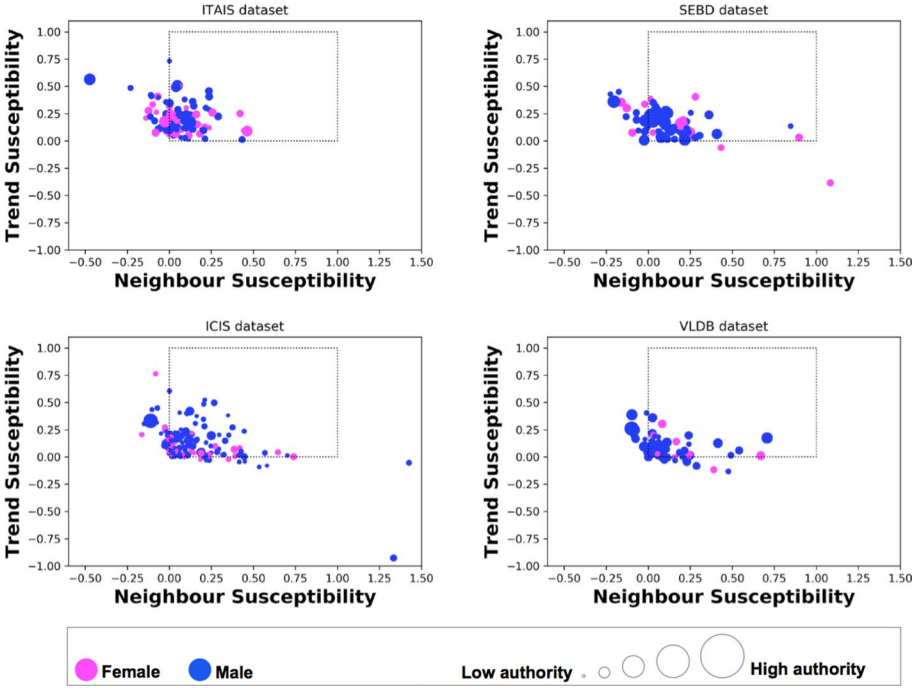

Fig. 12 Relationship between trend susceptibility, neighbour susceptibility and authority (ball size) for authors having authority greater than average

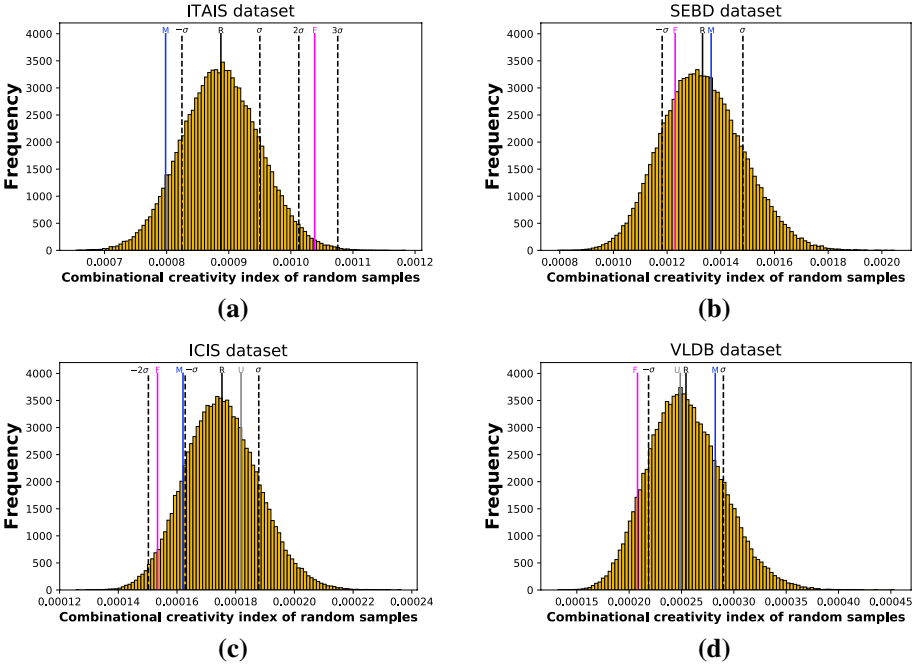

Fig. 13 Distributions of sample averages of the combinational creativity index. They refer, respectively, to **a** ITAIS, **b** SEBD, **c** ICIS, and **d** VLDB communities

Table 7 For each scientific community: $\bar{\mathcal{C}}_F$ is the average combinational creativity index value for females; $\bar{\mathcal{C}}_M$ is the average combinational creativity index value for males; $\bar{\mathcal{C}}_U$ is the average combinational creativity index value for undetermined authors; $\bar{\mathcal{C}}_R$ is the average combinational creativity index value for the random samples; σ is the standard deviation of the different averages of the random samples; z_F is the Z-score of the standard normal distribution corresponding to the average combinational creativity index values for each female group; the *p-value* is the likelihood given by the complementary cumulative distribution function

Community	$\bar{\mathcal{C}}_F$	$\bar{\mathcal{C}}_M$	$\bar{\mathcal{C}}_U$	$\bar{\mathcal{C}}_R$	σ	z_F	*p-value*
ITAIS	0.0010	0.0008	–	0.0009	6.27e−05	2.41	0.80%
SEBD	0.00123	0.00136	–	0.00133	0.000150	− 0.68	24.83%
ICIS	0.000153	0.000162	0.000182	0.000175	1.25e−05	− 1.75	4.01%
VLDB	0.000208	0.000282	0.000249	0.000254	3.57e−05	− 1.30	9.68%

Algorithm Algorithm for generation of random samples in pseudo-code.

```
k = 100000
m = sampleSize
n = numberOfAuthors
for all j ∈ [0,k) do
    randomSampleElements = Set()
    i = 0
    while i < m do
        randomNumber = randrange(n)
        if (randomNumber ∉ randomSampleElements) then
            randomSampleElements.add(randomNumber)
            i = i +1
        end if
    end while
end for
```

To improve readability, dashed black lines at multiples of the standard deviation (σ) of the distributions are also shown in the figures and reported in the tables. Then, for each indicator, the Z-scores (z_F) (Bluman 2009; Stuart and Ord 1987) of the female groups in the different communities are reported. These are the normalized deviations with respect to the distribution of the random samples. It is noteworthy that the Z-score is the effect size (see Coe 2002), where the mean of the experimental group is the average of the values for the members belonging to the group of females; the mean of the control group is the average value (for that indicator) of the 100,000 randomly generated samples of the same dimension of the female group. In details, we computed the average of the values for each sample and, hence, the mean of the control group results from the average of all these 100,000 average values; and the standard deviation (σ) is the standard deviation of the distribution of the samples. The *p-value* (i.e. the complementary cumulative distribution function) represents the likelihood of having either a higher average indicator value, if $\bar{ind}_F > \bar{ind}_R$, or a lower one, if $\bar{ind}_F < \bar{ind}_R$. Therefore, it measures how significant the observed average value of the indicator for the female groups (or its estimate for the ICIS and VLDB communities) is. The lower are the p's, the more significant are the deviations.

In this work, the hypothesis of gender divide for an indicator was generally accepted against the null hypothesis if the effect size had been higher than $\sigma/2$. However, such hypothesis was considered highly significant if the effect size had been higher than $5/2 \cdot \sigma$,

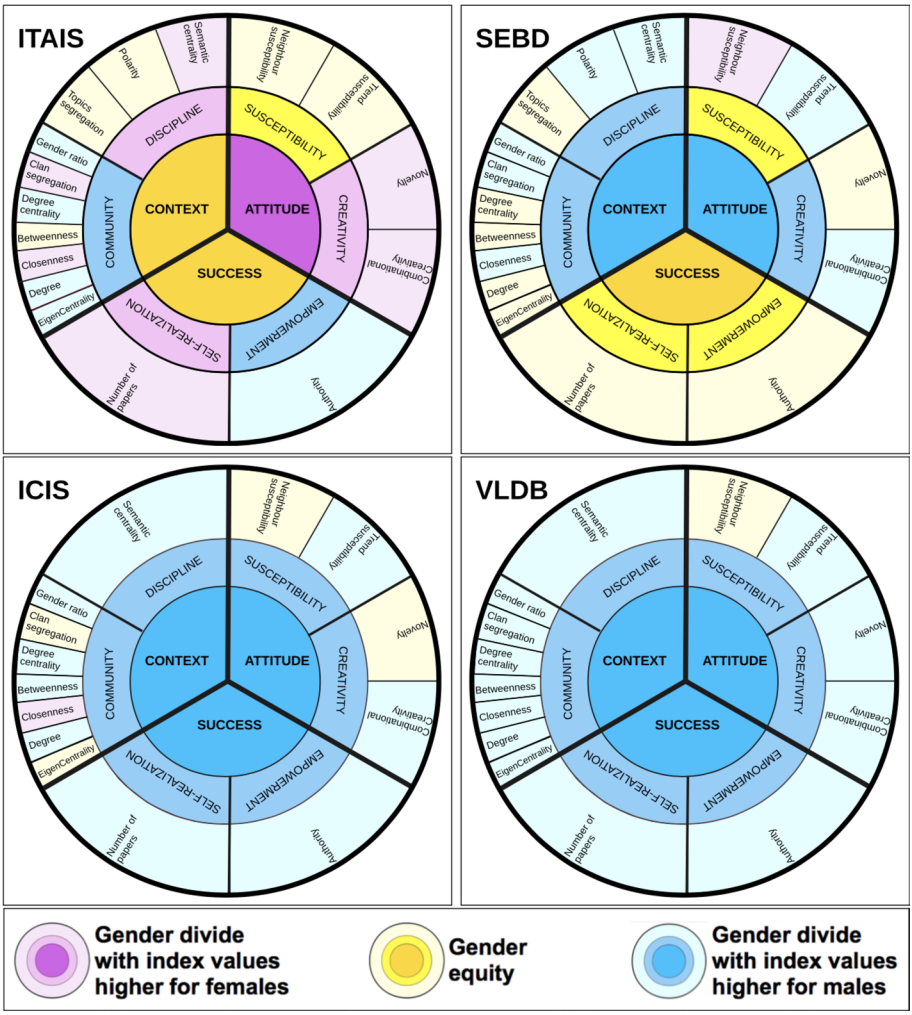

Fig. 14 Results of the assessment of the ITAIS, SEBD, ICIS, and VLDB communities by means of the gender divide framework

significant if the effect size had been between $3/2 \cdot \sigma$ and $5/2 \cdot \sigma$ and weakly significant if the effect size had been between $\sigma/2$ and $3/2 \cdot \sigma$. It was rejected if the effect size had been lower than $\sigma/2$. All such values are presented in Supplementary Information to the aim of full-reporting and transparency. Figure 14 provides a pictorial representation of the assessment framework (see also Fig. 1) that shows at glance the emerging gender divide in a community. In Fig. 14, shades of yellow color indicate gender equity, i.e. the observed gender divide is not significant from a statistical point of view (see the Supplementary Information), whereas shades of blue or pink colors indicate gender divide in case the mentioned hypothesis is considered, at least, weakly significant. Blue corresponds to the case the indicator value is higher for males and pink if it is higher for females. Table 8 presents a summary of the *p-values* for most of the indicators included in the gender diversity

Table 8 Summary table of *p-values*. Text formatted in bold or italics represents the case there is either an highly significant or a significant gender divide; the bold text represents the case the value of the indicator is higher for males, while the italic text when it is higher for females; and the normal text represents the case there is either a weakly significant or a not significant gender divide

ind_i	Description	*p-value* ITAIS (%)	*p-value* SEBD (%)	*p-value* ICIS (%)	*p-value* VLDB (%)
$\mathcal{SC}$	Semantic Centrality	20.34	15.87	15.15	10.20
$\mathcal{B}$	Betweenness	42.86	37.45	7.21	11.9
$\bar{D}_{cen}$	Degree Centrality	**0.16**	43.64	14.46	**4.27**
$\bar{Cl}$	Closeness	*2.02*	**1.92**	30.15	11.51
$\bar{E}$	EigenCentrality	**0.36**	32.64	33.72	14.23
$\mathcal{CS}$	Clan Segregation	*2.00*	**1.11**	40.50	12.36
$\bar{\mathcal{N}}$	Novelty	*4.09*	36.32	37.07	17.11
$\bar{\mathcal{C}}$	Combinational Creativity	*0.80*	24.83	**4.01**	9.68
$\bar{p}_{all}$	Number of papers	13.79	38.59	9.01	**6.43**
$\bar{D}_{all}$	Degree	**0.16**	43.64	14.46	**4.17**
$\bar{x}$	Neighbour susceptibility	33.72	*1.11*	38.59	39.36
$\bar{x}_s$	Trend susceptibility	31.56	11.51	**3.22**	20.32
$\bar{a}$	Authority	18.94	41.29	**6.18**	14.69

Table 9 Summary table of gender divide direction ($GD_{\nearrow}$) for polarity, topics segregation, and gender ratio. **M** represents the case there is gender divide and the indicator value is higher for males; *EQ* represents the case of gender equity; and **-** represents the case the indicator value is not available

ind_i	Description	$GD_{\nearrow}$ ITAIS	$GD_{\nearrow}$ SEBD	$GD_{\nearrow}$ ICIS	$GD_{\nearrow}$ VLDB
$\mathcal{P}$	Polarity	*EQ*	**M**	–	–
TS	Topics segregation	*EQ*	*EQ*	–	–
GR	Gender Ratio	**M**	**M**	**M**	**M**

framework for all the communities[9]. Finally, Table 9 contains a summary of gender divide direction ($GD_{\nearrow}$) for the other indicators where it was not possible to compute the *p-value*.

Discussion

In the following we answer to the 10 research questions presented in the "Introduction" section. We show how the gender divide framework can add something to the current discourse about gender and STEM. Since the answers to most of the research questions (from

[9] The presented analysis is not adjusted for multiple measures (e.g., see Bonferroni correction Bonferroni 1936).

RQ_2 to RQ_{10}) derive from the analysis of the four mentioned scientific communities, the validity of these conclusions is limited only to them.

(RQ_1) *Can we assess gender divide in a scientific community?*

Existing reports on gender diversity in STEM disciplines show that there are some differences in data related to males and females. Here, we considered communities of scientists as complex systems where the members interact each other and with the environment. Then, we proposed a framework to assess emergent gender diversity that addresses it from a multi-dimensional perspective including the context, the attitudes of the community members, and the achievements in terms of self-realization and empowerment. The objective of this framework is to be a tool to assess gender divide. The framework does not deal with the problem of determining if gender divide is due to individual features or to the surrounding environment. For this purpose, existing works suggest that such differences are mainly due to the latter. For instance, a recent study on gender gaps using data from third through eighth grade students in U.S. found that males are better than females in maths in socioeconomically advantaged school districts and in districts with larger gender disparities in adult socioeconomic status (Reardon et al. 2018).

We assessed four scientific communities arising from four conferences to test the gender divide framework. Two of them are based in Italy, while the other two are international. For both the Italian and the international communities, one community is related to the information system discipline, whereas the other one is related to computer science. From a general perspective, the cumulative analysis of the case studies provides evidence that the observed gender divide deals mainly with the indicators related to the context (23 out of 36 indicator values) and the success (5 out of 8 indicators) and less with the attitude (7 out of 16 indicator values).

(RQ_2) *As a group, are women and men equally successful?*

Considering the indicators related to the success dimension of the framework and, more in details, the number of papers and the authority, we observed that women and men are equally successful in the Italian communities, whereas men are more successful in the international ones.

(RQ_3) *As a group, are women and men equally creative?*

Considering the indicators related to the creativity metrics and, hence, the novelty and the combinational creativity, women are more creative than men in ITAIS, whereas men are more creative than women in VLDB. Finally, men tend to combine more existing concepts than women, whereas both the genders are equally capable in proposing novel concepts in SEBD and ICIS.

(RQ_4) *As a group, are women and men equally key to the larger community?*

Degree and degree centrality are two different indicators to measure the number of connections of a member of a community. Instead, betweenness, eigencentrality and closeness concern the probability that a new topic, for instance, reaches a member of a community by considering only the topological structure of the network. Hence, they also measure how key is the position of a member in the network in order to allow the diffusion of topics. The study reveals that there is a clear cut evidence that men are more key than women only in the VLDB community, whereas in the other communities there is more equity in that sense. Moreover, it is noteworthy that, even if men have more connections than women in ITAIS and ICIS (see degree and degree centrality), women have similar values of betweenness, eigencentrality and closeness and, hence, similar probability to diffuse topics. This means that women tend to connect more with key members with respect to what men do.

Semantic centrality measures to what extent the topics of a member belong to the mainstream and, hence, how key are the topics treated by her or him. Our study reveals that ITAIS is the only community where women treat more central topics, while men treat more central topics in the other ones. In general, men tend to be more central in the subject of their communities and, hence, more influenced by trends. They address popular topics, for instance, by combining existing ones.

(RQ_5) *As a group, are women and men equally important in determining the future direction of research?*

This question involves all the three dimensions of the framework since, to determine the future direction of research, a scientist should be capable of proposing novel ideas, be in a position of the social network that facilitates their sharing, and authoritative. In fact, her or his ideas should be accepted and, in case, re-elaborated by other scientists. The novelty index measures how much a scientist is proposing novel ideas, while, as mentioned, betweenness, eigencentrality and closeness measure the topological centrality of the scientist in the community and authority measures how much the ideas of an author are spread among her or his neighbours. Considering all these indices, men tend to determine the future direction of research more than women in VLDB. Indeed, this is the only community where all these indices are in favor of men. Then, our study reveals that, in ITAIS, even if women tend to be more creative than men and there is a substantial equity with respect to the topological indices, authority is in favor of men. In this regard, it seems that even if women have all the credentials to determine the future direction of research they still have to be recognized in that sense by the community. A similar phenomenon was observed also by a study on US PhD recipients by Hofstra et al. (2020) that revealed that women tend to introduce more novelty but have less impact. We observed gender equity for almost all these indices in SEBD and, hence, both the genders are equally important in determining the future direction of research in this community. Finally, authority and centrality indices are slightly in favor of men in ICIS whereas there is gender equity with respect to novelty. Again, also for this community, women have the potential to be important in determining the future direction of research.

(RQ_6) *As a group, do women and men study different topics?*

This question concerns the indices related to the discipline metrics and, hence, semantic centrality, polarity, and topics segregation. As mentioned, ITAIS is the only community where women study more central topics with respect to what men do. However, as shown by the values related to polarity of topics and topics segregation, our study did not reveal extreme cases of gender divide. Instead, even if the topics segregation issue was not observed in SEBD, women and men tend to address different topics. In particular, men tend to address more central topics in this community. Finally, we observed that men treat more central topics with respect to women also in the ICIS and VLDB communities.

(RQ_7) *As a group, do women and men work with each other?*

This question concerns three different indicators: clan segregation, neighbour susceptibility, and authority. Clan segregation indicates the presence in a community of prominent clusters of researchers consisting of either men or women. This indicator provides the analyst with a direct way based on the connections of the members to answer this question. Another less direct way to measure how effective is such collaboration is to look at the neighbour susceptibility and authority indices. In fact, they measure, respectively, how much a member of a community is influenced by her or his connections and how much a member influences the others.

Concerning clan segregation, ITAIS is the only community where this index is in favor of women. This means that women are more represented in the largest clans than

expected considering the gender ratio in the whole community. Clan segregation is in favor of men in ICIS and VLDB and there is equity in SEBD. Concerning neighbour susceptibility, there is equity in ITAIS, ICIS, and VLDB. Instead, in SEBD, women tend to be influenced by others more than men. In a certain sense, this is a positive aspect as it means that women "exploit" somehow their connections in the social network. Similarly, in SEBD all the authors have a similar capability of influencing the others, while men tend to influence more the others in ITAIS, ICIS, and VLDB. To conclude, even if from a topological perspective we observed that women in SEBD are less present in the largest clan, they are more likely to be influenced by others and are equally authoritative with respect to men. Hence, women of SEBD exploit their relationships more effectively than those of the other communities.

(RQ_8) *As a group, do women and men have different sources of inspiration?*

As mentioned above, SEBD is the only community where women are more influenced and, hence, inspired by others with respect to what happens to men. Men and women are equally influenced by their connections, even if neighbour susceptibility is slightly higher for women, in the other communities. This means that women tend to be influenced by their colleagues more than men. Both this tendency, the above-mentioned subject centrality of men and differences in gender ratios could be due, for instance, to the limited attendance of women to conferences where most of the novel ideas are spread in the community. This suggests the need of policies to increase such participation. Men are more influenced by trends in SEBD, ICIS, and VLDB. Instead, men and women are equally influenced by trends in ITAIS.

(RQ_9) *Does the gender divide depend on discipline?*

The assessment provides evidence that there are indicators that measure gender divide (21 out 30 in both the disciplines) in both CS and IS. However, while there are some indices (i.e. semantic centrality, closeness, clan segregation, novelty, combinational creativity, number of papers) that are in favor of women in IS, there are not indices in favor of women in CS, as the neighbour susceptibility index can not be categorized as in favor of one gender. Hence, gender divide in favor of men in CS is higher with respect to that in the IS discipline.

(RQ_{10}) *Within a discipline, does the gender divide vary on a geographical basis?*

Finally, the assessment of indicator values reveals that there is less gender divide in Italy with respect to the international communities. In fact, we observed gender divide in 19 out of 32 indicators in the Italian communities, whereas we observed gender divide in 22 out 28 indicators in the international ones. Moreover, there are 5 indicators that show gender divide in favor of women in the Italian communities, while only one indicator shows gender divide in favor of women in the international ones. Such results are in accordance with those presented for a related discipline in the global gender gap report (WEF 2018). In fact, such report reveals the outcome of an analysis of LinkedIn data showing that the gender gap concerning share of professionals with artificial intelligence skills in Italy is among the lowest. The presented data analysis does not motivate why the Italian case is different. This could be due to existing research policies or to the role of women in Italy. Differences from the international communities were observed but motivating them needs a deeper analysis of the Italian society.

Conclusions

The gender divide framework is a tool that allows to recognize gender divide in scientific communities and suggests some possible directions to intervene to reduce such issue. The above-described results provide insights on the assessed communities. For instance, the study reveals that females are more creative, write more papers than males and treat more central topics in ITAIS. However, in general, novelty is not always in favor of males or females. This demonstrates that the conjectures that females are not suited to these subjects are unfounded. The only indicator always in favor of males is gender ratio. This means that for both the information systems and computer science subjects, whatever the geographic context is, initiatives to increase the number of females should be taken. Other indicators, as centrality indicators, are less clear probably because there are not intrinsic differences in genders but the socio-cultural motivations play a relevant role.

The substantial conclusions are that research policy can do a lot in promoting effective female participation as females can provide a valuable support in the scientific advancement of computer science and information system disciplines. Then, the results suggest that gender divide depends on the discipline but the socio-cultural context plays a relevant role. Currently, there are several initiatives to this purpose that mainly aspire to increase participation of women. For instance, increasing participation of women in computer science is currently an objective of several prominent institutions as Carnegie Mellon University, Rice University, and of other initiatives as the AIS Women's Network (AISWN), the Anita Borg Institute, and the Women in Computational Intelligence sub-committee of the Computational Intelligence Society (Vardi 2015; Estevez 2017). Increasing representativeness of women in STEM could also break new ground in developing concepts, theories and methods (Fatourou et al. 2019). For instance, it could be possible to reduce the problem of having gender biased decision-making systems as those embedding the gender biased artificial intelligence algorithms described by Zou and Schiebinger (2018).

Appendix

According to the approach presented by D'Agostino et al. (2015), neighbour susceptibility (x_i) and trend susceptibility (x_{is}) for each member h_i are achieved analytically by solving the following equations:

$$\begin{pmatrix} \langle\langle [\Delta^N_{h_i}]^2 \rangle\rangle & \langle\langle \Delta^N_{h_i} \cdot \Delta^s_{h_i} \rangle\rangle \\ \langle\langle \Delta^N_{h_i} \cdot \Delta^s_{h_i} \rangle\rangle & \langle\langle [\Delta^s_{h_i}]^2 \rangle\rangle \end{pmatrix} \begin{pmatrix} x_i \\ x_{is} \end{pmatrix} = \begin{pmatrix} \langle\langle \delta\xi_{h_i}(c_k,t) \cdot \Delta^N_{h_i} \rangle\rangle \\ \langle\langle \delta\xi_{h_i}(c_k,t) \cdot \Delta^s_{h_i} \rangle\rangle \end{pmatrix} \Rightarrow$$

$$\mathcal{A} \cdot \begin{pmatrix} x_i \\ x_{is} \end{pmatrix} = \mathcal{B},$$

where $\Delta^s_{h_i} = [L_s(c_k,t) - L_{h_i}(c_k,t)]$ is the deviation of the semantic profile ($L_{h_i}(c_k,t)$) of a member h_i from the environment ($L_s(c_k,t)$); $\Delta^N_{h_i} = [L^N_{h_i}(c_k,t) - L_{h_i}(c_k,t)]$ is the deviation of the semantic profile of a member h_i from the average profile of her or his neighbours; , $\delta\xi_{h_i}(c_k,t) = \xi_{h_i}(c_k,t+\Delta t) - \xi_{h_i}(c_k,t)$ is the variation of the share of interest of a member h_i in the topic c_k; and $\langle\langle f \cdot g \rangle\rangle = \sum_{c_k,t} f(c_k,t)g(c_k,t)$ is the scalar product.

Hence:

$$x_i = \frac{\begin{vmatrix} \langle\langle \delta\xi_{h_i}(c_k,t)\cdot \Delta^N_{h_i}\rangle\rangle & \langle\langle \Delta^N_{h_i}\cdot \Delta^s_{h_i}\rangle\rangle \\ \langle\langle \delta\xi_{h_i}(c_k,t)\cdot \Delta^s_{h_i}\rangle\rangle & \langle\langle [\Delta^s_{h_i}]^2\rangle\rangle \end{vmatrix}}{\det \mathcal{A}} \tag{5}$$

$$x_{is} = \frac{\begin{vmatrix} \langle\langle [\Delta^N_{h_i}]^2\rangle\rangle & \langle\langle \delta\xi_{h_i}(c_k,t)\cdot \Delta^N_{h_i}\rangle\rangle \\ \langle\langle \Delta^N_{h_i}\cdot \Delta^s_{h_i}\rangle\rangle & \langle\langle \delta\xi_{h_i}(c_k,t)\cdot \Delta^s_{h_i}\rangle\rangle \end{vmatrix}}{\det \mathcal{A}} \tag{6}$$

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11192-021-03885-3) contains supplementary material, which is available to authorized users.

Acknowledgements Open Access funding provided by Ente per le Nuove Tecnologie, l'Energia e l'Ambiente. The authors kindly acknowledge Professor Cecilia Rossignoli, former President of the ITAIS Association, and the whole ITAIS community for fruitful discussion on gender diversity. They also thank Alida Zurke for a critical reading of the manuscript.

Data Availability The datasets generated and/or analyzed during the current study are available at https://tinyurl.com/genderdivide-dataset.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abraham, A. (2016). Gender and creativity: an overview of psychological and neuroscientific literature. *Brain Imaging and Behavior*, *10*(2), 609–618. https://doi.org/10.1007/s11682-015-9410-8.

Aral, S., & Walker, D. (2012). Identifying influential and susceptible members of social networks. *Science*, *337*(6092), 337–341. https://doi.org/10.1126/science.1215842.

Bavelas, A. (1950). Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America*, *22*(6), 725–730.

Bendels, M. H. K., Müller, R., Brueggmann, D., & Groneberg, D. A. (2018). Gender disparities in high-quality research revealed by nature index journals. *PLOS ONE*, *13*(1), 1–21. https://doi.org/10.1371/journal.pone.0189136.

Bluman, A. G. (2009). *Elementary Statistics: A Step by Step Approach*. New York: McGraw-Hill Higher Education.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, *424*(4–5), 175–308. https://doi.org/10.1016/j.physrep.2005.10.009.

Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, *92*(5), 1170–1182.

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del Regio Istituto Superiore di Scienze Economiche e Commericiali di Firenze*, *8*, 3–62.

Coe, R. (2002). It's the effect size, stupid: What effect size is and why it is important. In: Paper presented at the 2002 Annual Conference of the British Educational Research Association, University of Exeter, Exeter, Devon, England.

D'Agostino, G., D'Antonio, F., De Nicola, A., & Tucci, S. (2015). Interests diffusion in social networks. *Physica A: Statistical Mechanics and its Applications*, *436*, 443–461. https://doi.org/10.1016/j.physa.2015.05.062.

D'Agostino, G., & De Nicola, A. (2016). Interests diffusion on a semantic multiplex. *The European Physical Journal Special Topics*, *225*(10), 2033–2045. https://doi.org/10.1140/epjst/e2016-60028-2.

D'Agostino, G., & De Nicola, A. (2019). Analysis of gender diversity in the Italian community of Information Systems. In A. Lazazzara, R. C. D. Nacamulli, C. Rossignoli, & S. Za (Eds.), *Organizing for Digital Innovation: At the Interface Between Social Media, Human Behavior and Inclusion* (Vol. 27, pp. 97–109)., Lecture Notes in Information Systems and Organisation (LNISO) Cham: Springer International Publishing.

Damore, J. (2017). Google's ideological echo chamber: How bias clouds our thinking about diversity and inclusion. https://assets.documentcloud.org/documents/3914586/Googles-Ideological-Echo-Chamber.pdf. Accessed: 2019-04-26.

De Nicola, A., & Missikoff, M. (2016). A lightweight methodology for rapid ontology engineering. *Communications of the ACM*, *59*(3), 79–86.

Di Tommaso, G., Gatti, M., Iannotta, M., Mehra, A., Stilo, G., & Velardi, P. (2020). Gender, rank, and social networks on an enterprise social media platform. *Social Networks*, *62*, 58–67.

Di Tommaso, G., Stilo, G. & Velardi, P. (2017). A gendered analysis of leadership in enterprise social networks. In: Proceedings of the ICWSM Conference.

Duch, J., Zeng, X. H. T., Sales-Pardo, M., Radicchi, F., Otis, S., Woodruff, T. K., et al. (2012). The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PLOS ONE*, *7*(12), 1–11. https://doi.org/10.1371/journal.pone.0051332.

EC: Directorate-General for Justice and Consumers (European Commission): Report on equality between women and men in the EU 2018 (2018)

EIGE: European Institute of Gender Equality: Gender equality and youth: opportunities and risks of digitalisation - main report. https://eige.europa.eu/publications/gender-equality-and-youth-opportunities-and-risks-digitalisation (2019). https://doi.org/10.2839/148393. Accessed: 2019-04-09.

Elsevier: Gender in the global research landscape. Elsevier Report. Available: https://www.elsevier.com/research-intelligence/resource-library/gender-report (2017).

Estevez, P. A. (2017). Trusting machines and gender diversity [president's message]. *IEEE Computational Intelligence Magazine*, *12*(2), 3–17.

Fatourou, P., Papageorgiou, Y., & Petousi, V. (2019). Women are needed in STEM: European policies and incentives. *Communications of the ACM*, *62*(4), 52–57. https://doi.org/10.1145/3312565.

Ford, H. L., Brick, C., Blaufuss, K., & Dekens, P. S. (2018). Gender inequity in speaking opportunities at the american geophysical union fall meeting. *Nature Communications*, *9*(1), 1358.

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., et al. (2018). Science of science. *Science*, *359*, 6379. https://doi.org/10.1126/science.aao0185.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239.

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, *115*(16), E3635–E3644. https://doi.org/10.1073/pnas.1720347115.

Gero, J. S. (2000). Computational models of innovative and creative design processes. *Technological Forecasting and Social Change*, *64*(2), 183–196. https://doi.org/10.1016/S0040-1625(99)00105-5.

Gerow, A., Hu, Y., Boyd-Graber, J., Blei, D. M., & Evans, J. A. (2018). Measuring discursive influence across scholarship. *Proceedings of the National Academy of Sciences*, *115*(13), 3308–3313. https://doi.org/10.1073/pnas.1719792115.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, *102*(46), 16569–16572.

Hofstra, B., Kulkarni, V. V., Galvez, S. M. N., He, B., Jurafsky, D., & McFarland, D. A. (2020). The diversity-innovation paradox in science. *Proceedings of the National Academy of Sciences*, *117*(17), 9284–9291.

Holman, L., Stuart-Fox, D., & Hauser, C. E. (2018). The gender gap in science: How long until women are equally represented? *PLOS Biology*, *16*(4), 1–20. https://doi.org/10.1371/journal.pbio.2004956.

Huang, J., Gates, A. J., Sinatra, R., & Barabási, A. L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences*, *117*(9), 4609–4616.

Hughes, H. (2004). The Cambridge Encyclopedia of the English Language. *Reference Reviews*, *18*(3), 28–29.

Humbert, A. L. & Guenther, E. A. (2018). D3.2 Measuring gender diversity in research teams: methodological foundations of the Gender Diversity Index. GEDII EU Project deliverable.

Jadidi, M., Karimi, F., Lietz, H., & Wagner, C. (2018). Gender disparities in science? Dropout, productivity, collaborations and success of male and female computer scientists. *Advances in Complex Systems*, *21*(03n04), 1750011. https://doi.org/10.1142/S0219525917500114.

Ke, Q., Ferrara, E., Radicchi, F., & Flammini, A. (2015). Defining and identifying sleeping beauties in science. *Proceedings of the National Academy of Sciences*, *112*(24), 7426–7431. https://doi.org/10.1073/pnas.1424329112.

Kersey, A. J., Braham, E. J., Csumitta, K. D., Libertus, M. E., & Cantlon, J. F. (2018). No intrinsic gender differences in children's earliest numerical abilities. *npj Science of Learning*, *3*(1), 12. https://doi.org/10.1038/s41539-018-0028-7.

Kim, J. (2018). Evaluating author name disambiguation for digital libraries: a case of DBLP. *Scientometrics*, *116*(3), 1867–1886.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature News*, *504*(7479), 211.

Lerback, J., & Hanson, B. (2017). Journals invite too few women to referee. *Nature News*, *541*(7638), 455.

Leslie, S. J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*(6219), 262–265. https://doi.org/10.1126/science.1261375.

Lunnemann, P., Jensen, M. H., & Jauffred, L. (2019). Gender bias in Nobel prizes. *Palgrave Communications*, *5*(1), 46. https://doi.org/10.1057/s41599-019-0256-3.

Martin-Hernandez, J. (2013). Measuring robustness of complex networks. Ph.D. thesis, Technische Universiteit Delft.

Matsumoto, M., & Nishimura, T. (1998). Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, *8*(1), 3–30.

Mattauch, S., Lohmann, K., Hannig, F., Lohmann, D., & Teich, J. (2020). A bibliometric approach for detecting the gender gap in computer science. *Communications of the ACM*, *63*(5), 74–80.

Mika, P. (2007). Ontologies are us: A unified model of social networks and semantics. *Web Semantics: Science, Services and Agents on the World Wide Web*, *5*(1), 5–15.

Newman, M. E. J. (2010). *Networks: An introduction*. Oxford; New York: Oxford University Press.

Nielsen, M. W., Andersen, J. P., Schiebinger, L., & Schneider, J. W. (2017). One and a half million medical papers reveal a link between author gender and attention to gender and sex analysis. *Nature Human Behaviour*, *1*(11), 791–796. https://doi.org/10.1038/s41562-017-0235-x.

Nielsen, M. W., Bloch, C. W., & Schiebinger, L. (2018). Making gender diversity work for scientific discovery and innovation. *Nature Human Behaviour*, *2*(10), 726–734.

O'Dea, R. E., Lagisz, M., Jennions, M. D., & Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for STEM. *Nature Communications*, *9*(1), 3777. https://doi.org/10.1038/s41467-018-06292-0.

Payton, F. C., & Berki, E. (2019). Countering the negative image of women in computing. *Communications of the ACM*, *62*(5), 56–63. https://doi.org/10.1145/3319422.

Reardon, S., Fahle, E. M., Kalogrides, D., Podolsky, A., & Zárate, R. C. (2018). Gender achievement gaps in U.S. school districts. CEPA Working Paper No. 18-13. Retrieved from Stanford Center for Education Policy Analysis at http://cepa.stanford.edu/wp18-13 Last access 27 July, 2019.

Reilly, D., Neumann, D. L., & Andrews, G. (2019). Investigating gender differences in mathematics and science: Results from the 2011 trends in mathematics and science survey. *Research in Science Education*, *49*(1), 25–50. https://doi.org/10.1007/s11165-017-9630-6.

Sowa, J. F. (2006). Semantic networks. In *Encyclopedia of Cognitive Science*. Wiley. https://doi.org/10.1002/0470018860.s00065.

Sternberg, R. J., & Lubart, T. I. (1999). The concept of creativity: Prospects and paradigms. *Handbook of Creativity*, *1*, 3–15.

Stuart, A., & Ord, J. K. (1987). *Kendall's Advanced Theory of Statistics: Distribution theory; Vol. 2, Classical Inference and Relationship; Vol. 3, Design and Analysis, and Time-Series*. Granville: Charles Griffin.

UNESCO: Cracking the code: Girls' and women's education in science, technology, engineering and mathematics (STEM). Paris: UNESCO Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000253479 (accessed April 23, 2020) (2017).

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*(6157), 468–472. https://doi.org/10.1126/science.1240474.
Vardi, M. Y. (2015). What can be done about gender diversity in computing?: A lot!. *Communications of the ACM*, *58*(10), 5–5.
WEF: World Economic Forum: The global gender gap report 2018 (2018).
Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, *566*(7744), 378–382. https://doi.org/10.1038/s41586-019-0941-9.
Yang, Y., Chawla, N. V., & Uzzi, B. (2019). A network's gender composition and communication pattern predict women's leadership success. *Proceedings of the National Academy of Sciences*, *116*(6), 2033–2038. https://doi.org/10.1073/pnas.1721438116.
Zeng, X. H. T., Duch, J., Sales-Pardo, M., Moreira, J. A. G., Radicchi, F., Ribeiro, H. V., et al. (2016). Differences in collaboration patterns across discipline, career stage, and gender. *PLOS Biology*, *14*(11), 1–19. https://doi.org/10.1371/journal.pbio.1002573.
Zou, J., & Schiebinger, L. (2018). Design AI so that it's fair. *Nature*, *559*(7714), 324–326.
Zweben, S., & Bizot, B. (2018). *2017 CRA Taulbee Survey*. CRA: Computing Research News.