Scientometrics (2021) 126:9147–9183
https://doi.org/10.1007/s11192-021-04144-1

An analysis of top author citations in software engineering and a comparison with other fields

Kai Petersen[1,2] · **Nauman Bin Ali**[2]

Received: 16 March 2021 / Accepted: 20 August 2021 / Published online: 9 October 2021

Abstract

Ioannidis et al. provided a science-wide database of author citations. The data offers an opportunity to researchers in a field to compare the citation behavior of their field with others. In this paper, we conduct a systematic analysis of citations describing the situation in software engineering and compare it with the fields included in the data provided by Ioannidis et al. For comparison, we take the measures used by Ioannidis into consideration. We also report the top-scientists and investigate software engineering researchers' activities in other fields. The data was obtained and provided by Ioannidis et al. based on the Scopus database. Our method for analysis focuses on descriptive statistics. We compared software engineering with other fields and reported demographic information for the top authors. The analysis was done without any modifications to the ranking. In the later analysis, we observed that 37% of researchers listed as software engineers were not in the software engineering field. On the other hand, the database included a large portion of top authors (ca. 60% to 80%) identified in other software engineering rankings. Other fields using the database are advised to review the author lists for their fields. Our research's main risk was that researchers are listed that do not belong to our studied field.

Keywords Software engineering · Author citations · Self-citations · Comparison

Introduction

Scientific measures are collected on various levels, for example, to assess institutions (Karanatsiou et al. 2019), scientific publication forums (Pendlebury and Adams 2012), impact of particular publications (Garousi and Fernandes 2016), and individual researchers (Karanatsiou et al. 2019),(Karanatsiou et al. 2019). The rankings in software engineering investigated the contributions of researchers using different measures.

The most prominent ranking for software engineering researchers was done for a series of time intervals (e.g., Glass and Chen 2003 for the time period 1998–2002) and was published by the Journal of Systems and Software (JSS). The measure used was based on

✉ Kai Petersen
kai.petersen@bth.se

1 University of Applied Sciences, Flensburg, Germany

2 Blekinge Institute of Technology, Karlskrona, Sweden

publications in the most influential journals of the field. Also, depending on author configurations (number of authors and positioning), a journal publication was counted differently (e.g., single-author papers were weighted higher). For the most recent edition of the JSS ranking, the counting rules and measures were changed. For example, the new ranking included more publication forums, and author citations were considered. Another recent assessment of top authors in software engineering was presented by Fernandes (2014). Fernandes used two different measures to rank the researchers: fractional and harmonic authorship credit (cf. Hagan Hagen 2014). Just considering the software engineering field, different rankings use different measures and obtain the basis for ranking (selected publication forums) in different ways.

Ioannidis et al. (2020) provided a science-wide database containing the scientific measures for top authors. The measure used for the ranking was referred to as Comosite Citation index, which were introduced in an earlier publication by Ioannidis et al. (2016). The Composite Citations index combines different measures of author performance into a single measure. The database covers 21 research fields and also documents subfields within the fields (in total 174)[1]. Each researcher was classified according to fields and subfields. The database provides the opportunity to add to the previous rankings, as the same measures were used across different fields and sub-fields. Therefore, a comparison of fields and sub-fields becomes possible. Furthermore, previous software engineering rankings emerged from rankings done by other software engineering researchers. As a non-software engineering researcher constructed the database, the risk of unintentional biases is reduced.

This paper's main contribution is to provide top authors in software engineering based on the database provided by Ioannidis et al. More specifically, we make the following contributions:

- *C1: Compare software engineering with other fields in the database.* Given that Ioannidis et al. use the same measures across fields, a unique opportunity is provided, allowing for such a comparison.
- *C2: Characterize the top authors in software engineering based on the measures provided.* We investigate different measures, such as top countries in terms of the number of scholars ranked, top universities, as well as age profile and research productivity. We also investigate which other sub-fields (besides software engineering) are associated with the top authors.

It is essential to highlight that we first analyze the data as-is. That is, no modifications are done to the data-set not to introduce bias in the initial analysis.

After that, we review the data in detail concerning validity issues concerning the data for software engineers. We also compare the findings from analyzing the database with the results of previous software engineering rankings and reflect on the measures used. Consequently, our analysis provides insights to other fields that use the database by Ioannidis et al. for analysis of their fields.

Section 2 (Related work) presents previous rankings of software engineering authors. Section 3 describes the data collection and resulting database and the variables studied. Section 4 presents the results. In Sect. 5 we compare the rankings and discuss the measures used. Section 6 concludes the paper.

[1] The data file (Table-S8-Field-Subfield-Thresholds-career-2019.xlsx) can be found in Mendeley at https://dx.doi.org/10.17632/btchxktzyw

Related work

Several studies have used bibliometrics data to reflect on various aspects of software engineering (SE) research.

Mathew et al. (2019) used topic modeling and bibliometric data (author names, venues, and citations) to identify ten major topics in SE research. Their study showed the changes in the major topics over the years as well. They also confirmed the higher citation count of journal articles compared to conference papers. They also investigated whether there is a gender gap in the authorship of publications on the top topics.

Fernandes (2014) reviewed proceedings and volumes of 122 conferences and journals and identified that the number of co-authors in SE publications is increasing over the years. The results were re-validated in another study by Fernandes and Monteiro in another study (Fernandes and Monteiro 2017). The findings are useful as they highlight a positive trend indicating more collaborative research, which is necessary for dealing with complex real-world challenges. Furthermore, the results highlight the need for clear guidelines for being co-authors and describing the role and contribution of the co-authors in the study.

In a series of 15 articles published from 1994 to 2011, researchers have published a ranking of most productive scholars and institutions in systems and software engineering (Glass 1994; Wong et al. 2011). In these studies, they have relied on the number of papers published in a few key SE journals and conferences by authors in a sliding 5-year period.

Garousi and Fernandes (2016) identified the top 100 most cited papers in SE research among approximately 70,000 papers. They further noticed that about 43% of the papers had no citations at all (Garousi and Fernandes 2017). Garousi and Fernandes (2017) investigated if there are differences in citation patterns for paper publication venues and the authors' countries.

In our previous work, we have looked at citation behavior in SE literature. Molléri et al. (2018) analyzed if the number of citations to a publication is related to its reporting quality. They found some association between the reported rigor of the empirical study and the number of citations. Poulding et al. (2015) identified that the number of citations alone is insufficient for assessing the academic impact of a publication. They evaluated using a citation behavior taxonomy (Bornmann and Daniel 2008) to classify and analyze citations to highly cited papers in leading SE conference proceedings. In another study, Ali et al. (2020) similarly used another classification scheme (Teufel et al. 2006) to reflect on the nature and quality of citations.

To the best of our knowledge, the citation behavior in SE research and its comparison to other fields has not been investigated yet. This study fills this gap by examining and comparing the SE researchers' profile, academic impact, and engagement in multiple fields with other scientific fields.

Method

Our study was motivated by the possibility to see how distributions of citations differ between other fields and software engineering, which was enabled by the database provided by Ioannidis et al. (2020).

Research questions

We formulated three research questions, one question focusing on the comparative analysis (Contribution C1), and two focusing on a more detailed analysis of the top authors in software engineering (Contribution C2). We argue that it is interesting to look at this ranking compared to other rankings, as software engineering researchers have not collected the data, and the measures were also defined independently, which merits a new analysis of top scientists.

- *Research question 1 (RQ1, C1):* How do author citations in software engineering compare to other fields concerning the citation measures provided by Ioannidis et al.?
- *Research question 2 (RQ2, C2):* What are the demographics of the top authors (e.g., country of origin, institutions, etc.)?
- *Research question 3 (RQ3, C2):* In which other fields software engineering authors are active?

Data collection

Ioannidis et al. extracted the data using researcher profiles from the Scopus database. The extraction focused on author profiles. Overall a total of 7.92 million author profiles were extracted. The standardized citation data used in this study was created by Ioannidis et al. (2020) by collecting raw citation data from Scopus on May 6, 2020. Ioannidis et al. included the top 100k, and also authors were considered that belong to the top 2% in their subfield discipline, and otherwise would not have made it onto the list.

The focus was on career-long impact. That is, the study considered the statistics for the years 1960 to the end of 2019. The authors also continuously update the database for future analysis.

The ranking used the Composite Citation index, which is derived measure from multiple indicators, and described in the following section. It is also noteworthy that Ioannidis et al. removed self-citations for the ranking. They found that this impacted the ranking, as 4.9% of scientists in the list with self-citations are not on the list without self-citations.

Analysis

A range of measures has been presented based on each researcher's profile. Table 1 lists and defines the measures. While most measures are self-explanatory (e.g., np6019) or generally well known (e.g., h-index), Composite Citation index and the definition of fields and sub-fields are interesting to discuss in further depth.

Composite citation index: Ioannidis et al. (2016) recognizing the increasing trend of multi-authorship, argued for the need of a Composite Citation index that considers multiple indicators to assess a researcher's impact. They (Ioannidis et al. 2016) considered nine indicators but based on a correlation analysis selected the following six to comprise the Composite Citation index *c*:

nc: Total citations.

Table 1 Measures used in the analysis. The ID of the measures corresponds to the ID used in the paper by Ioannidis et al. (2020)

ID	Measure name	Description	Used to answer RQs
Scientometrics			
c (ns)	Composite citations	Composite measure taking various indices into consideration, used for ranking	RQ1, RQ2
np6019	Productivity	Number of papers produced in the time period 1960-2019	RQ1, RQ2
nc9619 (ns)	Citations	Total citations the author received in the time period 1960-2019	RQ1
h19 (ns)	h-Index	Hiersch-Index in 2019	RQ1
self%	Self citations	Percent of self-citations of overall citations received	RQ1
Demographics			
ca	Career age	Number of years active	RQ2
Country	Country	Country the researcher is located.	RQ2
Institution	Institution	Affiliation of the researcher	RQ2
field	field	One of 21 fields considered per researcher	RQ1,RQ2
Sub-field	Sub-field(s)	Two of 174 sub-fields considered per researcher (main sub-field and second sub-field)	RQ1, RQ2

h: Hirsch h-index, which is the maximum value h when h papers of an author have been cited at least h times.
hm: Schreiber Hm index, which provides co-author adjustment to h-index (Schreiber 2008).
ncs: total citations to sole-authored papers
ncsf: total citations to papers for which the researcher is the sole or the first author.
ncsfl: total citations to papers for which the researcher is the sole, first, or the last author.

The above six indicators are log-transformed to a value between 0 and 1. These values are summed together to derive the Composite Citation index C (Ioannidis et al. 2016).

To avoid introducing bias as software engineering researchers (analyzing the data for our field), we report top-scientists (top 2%) fully[2].

Comparative analysis (RQ1): To answer the first research question (RQ1), we used descriptive statistics. We compared the distribution of the different measures of software engineering to other fields. An alternative could have been to compare with all 174 sub-fields. However, this would have resulted in too many comparisons to comprehend the data easily. However, we analyzed the number of the 172 sub-fields being ranked above or below the median, minimum, and maximum values for software engineering.

Top software engineering authors (RQ2/RQ3): When analysing the top authors in software engineering, we used simple frequency analysis to rank the countries and institutions based on the number of top authors associated with them. The measures of career age and productivity are not included in the Composite Citation index. Thus, we investigated the association between the two measures. We also checked the correlation using Kendall's τ, which is more robust compared to other correlation measures when the data sets have outliers (Croux and Dehon 2010).

Results

The results present the answers to the research questions, namely comparison with other fields (RQ1) and after that the detailed analysis of top authors in software engineering (RQ2).

RQ1: comparison with other fields

As described earlier, the authors were assigned to fields and sub-fields. Concerning the field "Information & Communication Technologies (ICT)", we compared the ICT subfield software engineering with other subfields in ICT. Here, we focus on the comparison of the Composite Citation index. Later, we compare software engineering with the overall ICT category regarding the other measures (see Table 1).

[2] The standardized data by Ioannidis et al. can be found on Mendeley at https://dx.doi.org/10.17632/btchxktzyw

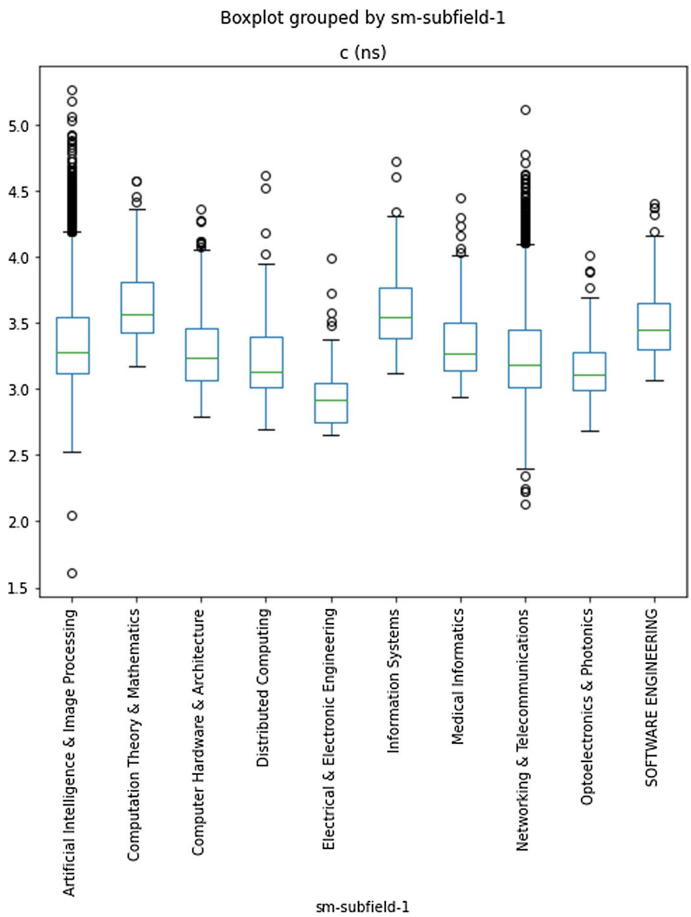


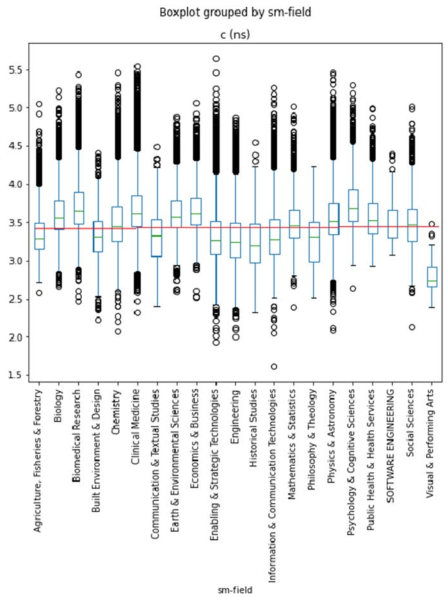
Fig. 1 Comparison of software engineering with sub-fields in the field "Information & Communication Technologies"

Comparison of software engineering with other ICT subfields

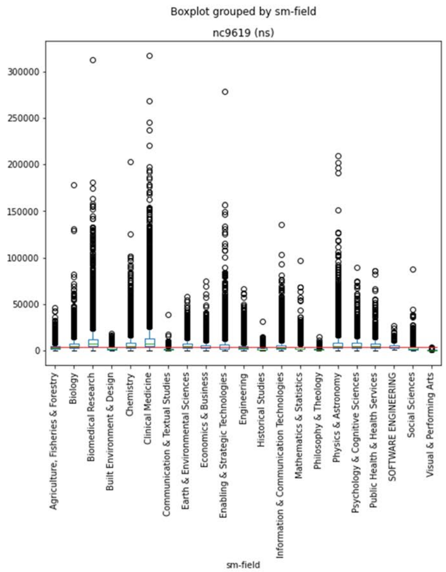
Figure 1 shows the distribution of the Composite Citation index of software engineering (SE) and other subfields in the ICT category. Only two fields have higher median values than SE, namely "Computational Theory & Mathematics" and "Information Systems". It is also noteworthy that the fields "Networking & Telecommunications" and "Artificial Intelligence & Image Processing" have a higher number data points (researchers) that are outliers concerning their Composite Citation index.

Comparison of software engineering with other fields

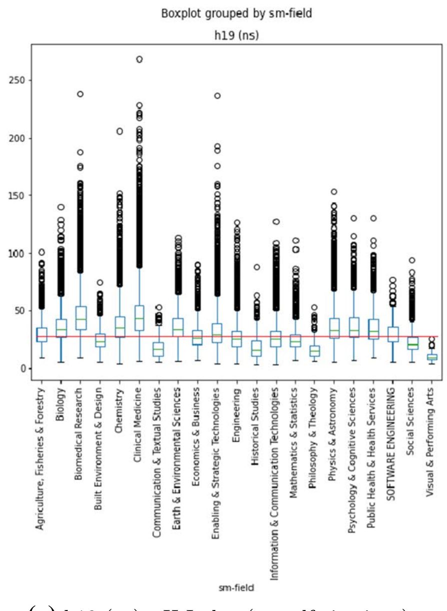
Figure 2 shows the comparison of software engineering with other fields concerning the distribution of the studied measures (see Table 1). Table 2 shows the ranking of software engineering and the other fields based on medians.



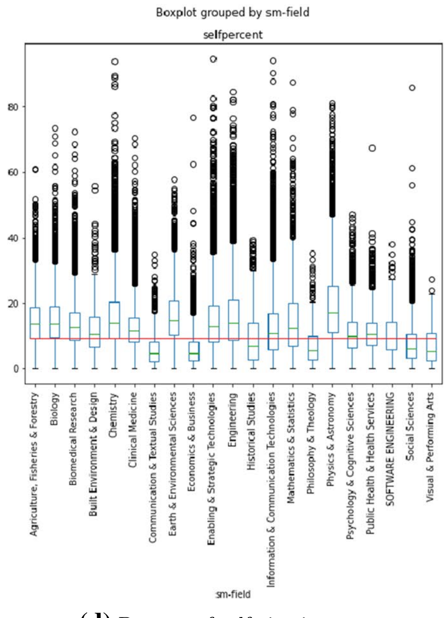
(a) c (ns) - Composite Citation index



(b) nc9619 (ns) - Citations 1960-2019 (no self-citations)

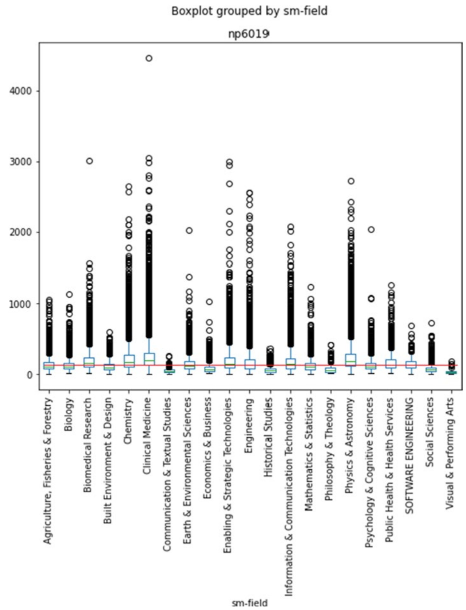


(c) h19 (ns) - H-Index (no self-citations)



(d) Percent of self-citations

Fig. 2 Data distributions for software engineering and comparison with other fields



(e) np6019 - Productivity 1960-2019

Fig. 2 (continued)

Composite Citation index (c): As described earlier (see Sect. 3) the Composite Citation index is a derived measure taking several indices into consideration. Software Engineering is ranked 9th (see Table 2). Seventeen of the considered fields have higher Composite Citation index values than the top Software Engineering researcher.

Citations (np9619): Concerning total cites, Software Engineering resides on rank 10 with a median of 3600 citations. Two fields stand out with more than twice the citations of Software Engineering, namely Clinical Medicine and Biomedical Research, both with more than 7000 citations (median). When looking at outliers (Fig. 2b), we see that 16 fields have top scientists with higher cites than the most highly cited researcher in Software Engineering. Four fields stand out when looking at the outliers, namely Clinical Medicine, Biomedical Research, Engineering & Strategic Technologies, as well as Physics & Astronomy.

H-Index (h19): Regarding the H-Index, Software Engineering (median of 28) is also ranked in the center of the fields considered (see Fig. 2c). Here, the fields of clinical medicine and biomedical research stand out with median H-Indexes of 43 and 42, respectively. Inspecting the data distributions and outliers, we can see that 16 of the fields have a higher H-Index as the highest-rated Software Engineering researcher.

Percent of self-citations (self%): The final measure considered is the percent of self-citations. One should note that the authors' ranking reported the statistics with and without self-citations. As pointed out by Ioannidis et al., the ranking changed by excluding self-citations. Authors who would have been in the top-list, considering self-citations, are not on the list with removed self-citations. Software Engineering is ranked 16th concerning the % of self-citations with a median value of 9.28%. No Software Engineering Author (in Ioannidis et al. dataset) has a self-citation ratio higher than 40%. On the other hand, we find several researchers with self-citations higher than 50% in other fields. Among the 159683 top authors (all fields), 627 have more than 50% of their citations as self-citations.

Table 2 Software engineering and other fields ranked by medium values

Rank	np6019		nc9619 (ns)		h19 (ns)		c (ns)		selfpercent	
	field	Medians	field	Medians	field	Medians	field	Medians	field	Medians
1	Clinical medicine	201.0	Clinical medicine	7426.0	Clinical medicine	43	Psychology and cognitive sciences	3.68	Physics and astronomy	17.05
2	Physics and astronomy	184.0	Biomedical research	7125.0	Biomedical research	42	Biomedical Research	3.65	Earth and environmental sciences	14.96
3	Chemistry	178.0	Psychology and cognitive sciences	4798.0	Chemistry	35	Economics business	3.61	Chemistry	14.10
4	Biomedical research	159.0	Physics and astronomy	4797.5	Biology	34	Clinical medicine	3.61	Engineering	14.01
5	Enablic strategic technologies	154.0	Chemistry	4782.0	Earth and environmental sciences	34	Earth and environmental sciences	3.58	Biology	13.83
6	ICT[1]	143.0	Earth and environmental sciences	4644.0	Physics and astronomy	33	Biology	3.56	Agriculture, fisheries and Forestry	13.72
7	Public health and health services	140.0	Biology	4503.0	Psychology and cognitive sciences	33	Public health and health services	3.52	Enablic strategic technologies	13.02
8	Engineering	137.0	Public health and health services	4193.5	Public health and health services	32	Physics and astronomy	3.51	Biomedical research	12.63
9	Software engineering	130.0	Economics Business	3652.0	Enablic strategic technologies	29	Social sciences	3.47	Mathematics and statistics	12.41
10	Earth and environmental sciences	124.0	Software engineering	3600.0	Agriculture, fisheries and forestry	28	Mathematics and statistics	3.46	Clinical medicine	11.54
11	Psychology and cognitive sciences	114.0	Enablic strategic technologies	3460.0	Software engineering	28	Software engineering	3.45	ICT	10.74
12	Agriculture, fisheries and forestry	113.0	ICT	3030.5	Economics business	26	Chemistry	3.44	Built environment design	10.58
13	Biology	113.0	Agriculture, fisheries and forestry	2815.5	Engineering	25	Communication and textual studies	3.33	Public health and health services	10.47
14	Mathematics and statistics	104.0	Engineering	2453.5	ICT	25	Built Environment design	3.32	Psychology and cognitive sciences	9.84

Table 2 (continued)

Rank	np6019		nc9619 (ns)		h19 (ns)		c (ns)		selfpercent	
	field	Medians	field	Medians	field	Medians	field	Medians	field	Medians
15	Built environment design	96.0	Mathematics and statistics	2280.0	Built Environment design	23	Philosophy and theology	3.31	Software engineering	9.28
16	Economics business	73.0	Social sciences	2191.0	Mathematics and statistics	23	Agriculture, fisheries and forestry	3.29	Historical studies	6.85
17	Social sciences	67.0	Built environment design	2124.0	Social sciences	21	ICT	3.28	Social sciences	6.12
18	Philosophy and theology	57.0	Communication and textual studies	1401.0	Communication and textual studies	17	Enablic strategic technologies	3.27	Philosophy and theology	5.42
19	Historical studies	53.0	Historical studies	1208.0	Historical studies	16	Engineering	3.24	Visual and Performing Arts	5.32
20	Communication and textual studies	47.0	Philosophy and theology	1127.0	Philosophy and theology	15	Historical studies	3.20	Communication and textual studies	4.79
21	Visual and performing arts	29.5	Visual and performing arts	411.0	Visual and performing arts	9	Visual and performing arts	2.74	Economics business	4.78

[1] Information and communications technology.

Table 3 Number of sub-fields with medians, minimum and maximum values lesser and greater than for software engineering

Measure ID	Measure	Median		Min		Max	
		< S	≥ S	< S	≥ S	< S	≥ S
c (ns)	Composite cit.	80	93	74	99	130	43
np6019	Productivity	89	84	30	143	64	109
nc9619 (ns)	Citations	91	82	93	80	55	118
h19 (ns)	h-Index	77	96	14	159	69	104
self\%	Self citations	51	122	0	173	58	115

Productivity (np6019): The number of total papers indicates the top authors' productivity. Software Engineering is on rank 9 of 21 with median productivity of 130 papers (Table 2). Fields such as clinical medicine, physics and Astronomy, and chemistry show substantially higher values. Clinical medicine has the highest productivity with 201 papers. The data distributions visualize the outliers among the top authors with very high productivity. While the top software engineering author has less than 100 publications, a substantial number of authors in the other fields have authors with substantially more publications. For example, in the areas Clinical Medicine, Chemistry, Engineering, and Physics & Astronomy, we find authors with more than 2000 publications.

Comparison with sub-fields: Given that each field has a number of sub-fields, we also compared software engineering with the data of the individual sub-fields. As the number of sub-fields is high, we could not show the distribution of the data side-by-side. Instead, Table 3 indicates the number of sub-fields with equal or greater numbers and lesser numbers when compared with software engineering. The medians for software engineering are in the center of the median values data set except for self-citations. There, 122 sub-fields have higher medians. This was already evident when looking at the aggregated data on the field level (see Fig. 2d. For minimum values, software engineering is also in the center of the data set for the Composite Citation index and total citations. However, for productivity and h-index, most sub-fields have higher minimum-values. When looking at the maximum values for the Composite Citation index, the majority of sub-fields are below software engineering. However, when just looking at citations, the majority is above. For productivity, h-index, and self-citation ratio, also the majority of sub-fields are above.

Overall, the key findings of the comparison of Software Engineering with other fields can be summarized as follows:

- Within related nine sub-fields in the field Information & Communication Technologies, Software Engineering is ranked third.
- When comparing Software Engineering with other fields, Software Engineering, lies in the center when looking at median values for the measures Total Papers (Productivity), Total Cites, H-Index, and Composite Citation index.
- Comparing the outliers (top of the top researchers) Software Engineering has considerably lower values than other fields. At least 15 or more fields have higher values than the best performing Software Engineer for the measures Total Papers (Productivity), Total Cites, H-Index, and Composite Citation index.
- In comparison to other fields, Software Engineering has a relatively low self-citation ratio of 9.28%.

Table 4 Top authors counts in software engineering per country

Group ID	Count	Countries
1	235	USA
2	32	Canada
3	28	UK
4	26	Germany
5	12	France, Switzerland
6	11	Italy
7	9	China
8	8	Australia, Israel
9	7	Sweden
10	6	Austria, Netherlands
11	5	Norway
12	4	Japan
13	3	Korea, New Zealand, Spain
14	2	Belgium, Ireland, Singapore
15	1	Czechia, Finland, Greece, Hong Kong, India, Luxembourg, Malta, Saudi Arabia, South Africa, Taiwan

RQ2: top-scientists

In this section, we present the top authors in software engineering and conduct after that characterize the ranked authors analyzing the measures provided by Ioannidis et al. Specifically, we first report countries and institutions. After that, as Career Age and Total Papers are not considered in the Composite Citation index, we investigate the associations between Composite Citation index and Career Age and Total Papers, respectively.

Top authors: Table 8 shows the complete list of the top 2% authors in Software Engineering. The total number of top authors is 441.

Countries: Table 4 lists the number of top authors per country. The USA stands out with the number of top-ranked scientists. Canada, UK, Germany, France, Switzerland, and Italy have more than ten scientists on the top list. It is also noteworthy that smaller countries (population with less than 10 million persons) are present who have multiple top authors, such as Switzerland (8.5 million), Israel (9.1 million), Sweden (10.2 million), Austria (8.9 million), and Norway (5.3 million).

Institutions: Table 5 shows the number of top scholars per institution. The last known institution/affiliation was used. The majority of institutions listed (172) only have one scholar working with them. We see that both universities and companies have active and hence recognized researchers in the top author list. The company with the most scholars is Microsoft Research (15 scholars), followed by Google LLC (9 scholars). Other companies highlighted are NVIDIA (7 scholars), Adobe Inc. (4 scholars), Pixar Animation Studios (4 scholars), and Facebook Inc. (2 scholars). The universities with the most scholars are the University of Washington, Carnegie Mellon University, Georgia Institute of Technology, with more than seven scholars each, and Stanford University, The University of British Columbia, and University of California, Davis.

Table 5 The number of top scholars per institution

ID	#Top scholar	#Inst.	Institution names
1	15	1	Microsoft Research
2	9	2	Google LLC, University of Washington
3	8	1	Carnegie Mellon University
4	7	1	Georgia Institute of Technology, NVIDIA
5	6	3	Stanford University, The University of British Columbia, University of California, Davis
6	5	3	Princeton University, UCL, University of Maryland, University of California, Irvine
7	4	13	Adobe Inc., Brown University, Delft University of Technology, Ecole Polytechnique Fédérale de Lausanne, INRIA Institut National de Recherche en Informatique et en Automatique, MIT Computer Science & Artificial Intelligence Laboratory, NC State University, Pixar Animation Studios, Politecnico di Milano, Rice University, The University of Texas at Austin, University of California, Berkeley, University of Massachusetts Amherst
8	3	16	Arizona State University, Blekinge Tekniska Högskola, California Institute of Technology, Colorado State University, George Mason University, Fairfax Campus, New York University, Technion - Israel Institute of Technology, The University of North Carolina at Chapel Hill, The University of Utah, University of Alberta, University of California, San Diego, University of Illinois at Urbana-Champaign, University of Southern California, University of Victoria, Universität Stuttgart, Victoria University of Wellington
9	2	39	Chalmers University of Technology, Cornell University, ETH Zürich, Facebook, Inc., Hebrew University of Jerusalem, Imperial College London, Istituto di Scienza e Tecnologie dell'Informazione A. Faedo, Johannes Kepler Universitat Linz, Monash University, National Institute of Standards and Technology, Naval Postgraduate School, Northeastern University, Open University, Oregon State University, Peking University, Purdue University, Queen's University, Kingston, Stony Brook University, Technical University of Berlin, Technische Universitat Wien, The Ohio State University, Universitetet i Oslo, University of Calgary, University of California, Los Angeles, University of Connecticut, University of Pennsylvania, University of Pittsburgh, University of Sheffield, University of Texas at Dallas, University of Toronto, University of Virginia, University of Wisconsin-Madison, University of Zurich, Università degli Studi del Sannio, Università della Svizzera italiana, Universität Konstanz, Vanderbilt University, Virginia Polytechnic Institute and State University
10	1	172	See Iaonnes et al. database for full list.

Fig. 3 Total papers (np6019) versus composite citation index

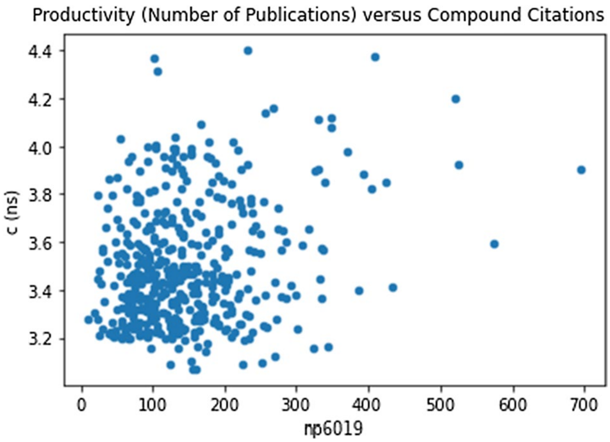
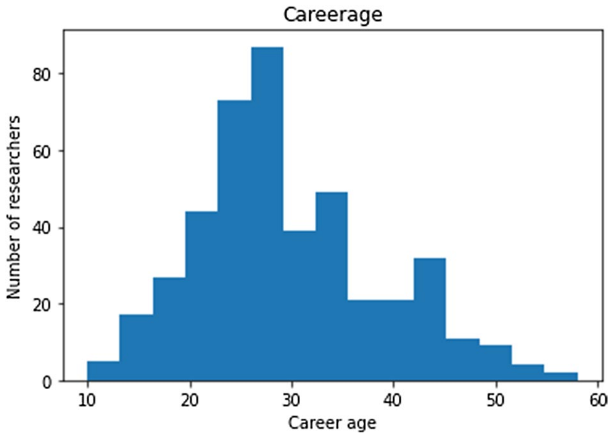


Fig. 4 Distribution of career age



Productivity and Composite Citation index: Figure 3 shows a scatter plot of the total papers produced in the time period 1960–2019 (np6019) and Composite Citation index (c(ns)). Visually the figure shows that there is no or little apparent association between the total number of papers produced and the Composite Citation index. Calculating the correlation between the two variables, we also see a very weak positive correlation (Kendall's $\tau = 0.091$).

Career age and composite citation index: Figure 4 shows the age profile of the career age of the top software engineering researchers. The figure shows that the majority of researchers have a career age of 25–35. However, when looking at the association between career age and Composite Citation index, we observe a similar pattern for productivity (papers published), i.e., the data lack association. This lack of association is also evident from the weak correlation (Kendall's Tau = 0.102). However, this may also be an artefact, given that the age of publications is not accounted for in the analysis (Fig. 5).

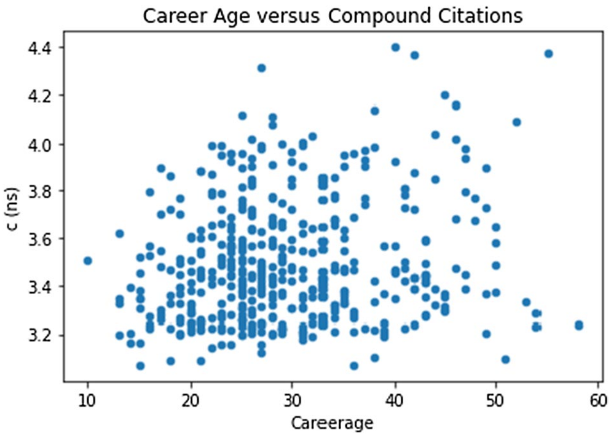


Fig. 5 Career age versus composite citation index

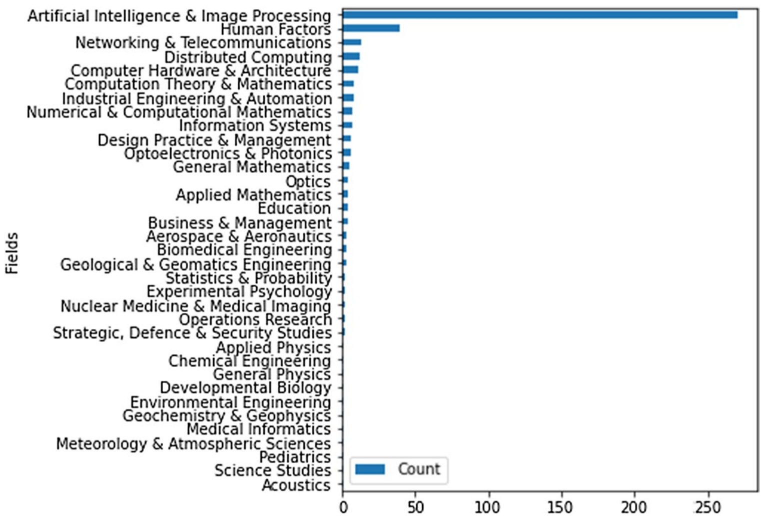


Fig. 6 Frequency of related fields occurring for top-authors

RQ3: Activity of top authors in other fields

As described earlier (see Table 1), in the data provided by Ioannidis et al., the author profiles were assigned two fields, one being the primary and the other the secondary sub-field. Figure 6 shows the count of secondary sub-fields for software engineering researchers.

Artificial Intelligence & Image Processing stands out as the secondary sub-field occurring most of the time. The second most frequently identified sub-field is Human Factors. After that, several sub-fields related to other computing areas (such as Distributed Computing or Computer Hardware & Architecture) are next based on counts. Fields further away from software engineering (e.g., General Physics, Development Biology, or Acoustics) are only occurring once.

Concerning the top authors as identified by Ioannidis et al., we highlight the following findings:

Table 6 Assessment of accurate classification of researchers in the top-list

	Total	Percent
In SE	278	63
Not in SE	163	37
	441	100

- Overall, 411 top authors were identified. The top three authors are Terzopoulos, Demetri, Boehm, Barry, and Levoy, Marc.
- The top three countries with the most top authors are the USA, Canada, and the UK.
- Concerning institutions, companies as well as universities are at the top. Microsoft leads with the highest number of top scholars (15), followed by Google LLC and the University of Washington.
- Productivity and career age are not associated with Composite Citation index, as indicated by correlation analysis.
- Artificial Intelligence & Image Processing is the most strongly associated second field for researchers with software engineering as their main field.

Discussion

When presenting the results, we kept the data as is, without filtering or manipulating the data. The reason was not to introduce biases in the analysis. After that, we reviewed the data set's details to reflect on the set for our particular case, i.e., software engineering.

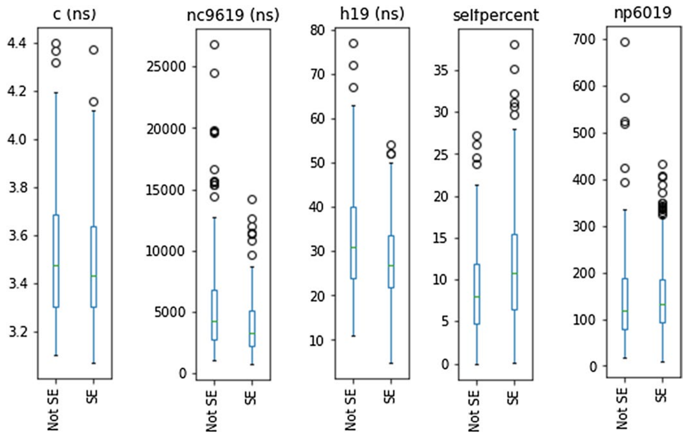
That is, in the discussion, we focus on two aspects when analyzing the results. First, we reviewed the top researchers' list to identify whether all are indeed in software engineering, given that the fields and sub-fields have been assigned using a trained machine-learning model. Second, we compared our findings with the rankings presented in the related work.

Reflections on the top-list

Looking at the authors, we classified them to either be in software engineering or not software engineering using Google Scholar profiles. If no profile was available, we identified the researchers' webpages. We looked at the research topics specified by the authors and their publication lists. Table 6 shows the analysis outcome. The focus was on identifying the researchers' main area, given that the main sub-field was under consideration. If software engineering is not the researchers' main area, this does not mean that the researchers classified with "Not in SE" have not done any software engineering research.

Thus, this is an indication that the machine-learning classifier was not perfect. In this step, we only detected that people with software engineering not being their main field appeared in the software engineering list.

For example, when looking at the top-list (see Table 8 in the Appendix), we see that among the top researchers, we have three cases of authors (Terzopoulos, Demetri; Levoy, Marc; Hoppe, Hugues) working in computer graphics. They all were classified in their second field as working with Artificial Intelligence & Image Processing. Thus, it appears that in some cases, the priorities in the fields (main field and second field) were not correct.



(a) c **(b)** nc9619 **(c)** h19 **(d)** Self% **(e)** np6019

Fig. 7 Comparison of true positives (software engineering researchers) and false positives (non-software engineering researchers) in the software engineering data set

When one analyzes the top researchers in computer graphics, these researchers would be missing in that analysis when looking at the main field.

However, it is essential to point out that many researchers were to be analyzed, requiring an automatic classification approach. Reviewing the data, we suspect that the classifier may have associated graphics and visualization in general with the strong focus of software engineering on end-user experience and usability and software visualization (e.g., researcher Eick, Stephen G.) as almost all researchers classified wrongly were associated with image processing, computer graphics, or visualization. Note that this is an interpretation observing the fields, though, given the nature of the model (convolutional neural networks) used for classification, it is retrospectively impossible to explain why the classifier produced the false positives.

Next, we analyze the impact on our analysis by comparing the SE-researchers with the non-SE researchers. In particular, we briefly describe how software engineering data was impacted by the inclusion of false positives (i.e., researchers classified as software engineers but should not have been). Figure 7 shows that the false positives impacted the software engineering data set. In particular, for Composite Citation index (c), total cites (nc9619), and h-index (h19), the values are moderately higher compared to the true positives. We also see that the extreme values (outliers) are higher for c, nc6919, and h19. For self-citations and total papers, the median values are slightly higher for the true positives.

When looking at the impact of the new median values on the ranking compared to other fields, the ranking in Table 2 would change. For Composite Citation index (new median = 3.432) and h-Index (new median = 27), the rank would be reduced by one. For Total Citations (np6919, new median = 3302), the rank would be reduced by two. Concerning Total Papers (np6019, new median = 133), there is no change in rank. However, for Percent of Self-Citations, the largest change occurs, increasing software engineering's rank by five (new median = 10.75%). In Table 8 in Appendix A we highlighted the correctly classified software engineering researchers using (∗).

Next, we compare the ranking based on the database provided by Ioannidis et al. with the already published rankings of software engineering scholars.

Comparison with related work

Two prominent rankings focused on identifying the top scholars in software engineering. The first is the JSS-Ranking, which was published for different time intervals (see Sect. 2). Furthermore, Fernandes (2014) published a ranking covering the years 1971–2012.

JSS-rankings: The first JSS-ranking has been published in 1994 (Glass 1994). Thereafter, each year the ranking was extended by 1 year (e.g., 1993–1994 Glass 1995 and thereafter 1993–1995 Glass 1996). Later, not only the last year was incremented, but also the first year (see e.g. rankings 1993–1997 Glass 1998 and 1994–1998 Glass 1999. The last ranking using consistent measures and schedules was published in 2011 (Wong et al. 2011). A new ranking was published after a break, using new scientific measures (Karanatsiou et al. 2019). To compare the findings, we first compare the series from 1993 to 2008 with our findings and compare with the latest ranking (cf. Karanatsiou et al. 2019).

Regarding the rankings from 1993 to 2008, the rankings were only based on the journals that were considered the top ones in software engineering. Each paper published by an author was counted differently depending on the number of co-authors. A single author would receive one point for a paper. For co-authors, first fractional contributions were calculated (e.g., having three authors would give 0.33 points for each author). The fractional values were transformed (e.g., a fractional contribution of 0.33 would become 0.5) to not penalize co-authorship. Given that the ranking is purely focused on productivity (number of papers published), we should not expect that most JSS ranking authors appear in the database by Ioannidis, where the focus was on a Composite Citation index (c). However, we expect some overlap and may see whether researchers appearing in the list are in the list by Ioannidis et al., but not classified as software engineers.

Table 7 provides an overview of the ranking results, focusing on the rankings for the time periods 1993–1997 (Glass 1998) until the last ranking (Wong et al. 2011), which included the time periods 2003-2007 as well as 2004-2008. The table shows the ranks the researchers achieved in a particular ranking. In case an author did not appear in the ranking, this is indicated using the bullet-symbol (•). In the outer right column, we state whether the authors in the JSS ranking appeared in the ranking by Ioannidis et al. If a rank-number is stated, they appeared in the software engineers' list, whose fields were correctly classified as software engineering. In case the author did not appear, we used the bullet-symbol again (•).

We also show which authors were found in the top-list we generated based on the database by Ioannidis et al.

We distinguish between finding the authors among those classified as software engineering researchers in their main field (the rank number is stated), or researchers not classified as software engineers in their main field, but included and not classified as software engineers in the database ($\surd$). As is shown in Table 7 21 of the 54 (39%) researchers that were classified by Ioannidis et al. as being in software engineering (primary field)) were also ranked by JSS. We also found six researchers (11%) in the JSS ranking but not classified with software engineering as their primary field in the database.

Table 7 Overview of the JSS ranking and the researchers included in the Ioannidis et al. database

No.	Name	Rankings												
		(Glass 1998)	(Glass 1999)	(Glass 2000)	(Glass and Chen 2001)	(Glass and Chen 2002)	(Glass and Chen 2003)	(Glass and Chen 2005)	(Tse et al. 2006)	(Eric Wong et al. 2008)	(Eric Wong et al. 2009)	(Wong et al. 2011)	(Wong et al. 2011)	Ioannidis et al.
1	Chatzigeorgiou, Alexander	•	•	•	•	•	•	•	•	•	•	15	•	•
2	Angelis, Lefteris	•	•	•	•	•	•	•	•	•	•	•	4	•
3	Basili, Victor R.	4	4	4	2	7	12	•	•	•	•	•	•	6
4	Bieman, Jim	•	12	14	•	•	•	•	•	•	•	•	•	312
5	Bosch, Jan	•	•	•	•	•	•	•	•	13	15	•	•	55
6	Briand,Lionel	•	•	•	•	•	9	5	5	4	4	5	5	11
7	Chang, Chin-Chen	•	•	10	8	10	8	•	•	•	15	•	8	√
8	Chang, Ruei-Chuan	•	•	7	8	7	•	•	•	•	•	•	•	•
9	Chen, Jen-Yen (Jason)	6	8	•	14	•	•	•	•	•	•	•	•	•
10	Chen, Tsong Yueh	•	5	10	12	9	7	12	6	6	•	•	•	233
11	Chou, Shih-Chien	•	•	•	•	•	•	•	13	4	3	2	2	•
12	Chung, Chin-Wan	•	•	•	•	•	•	•	•	8	7	13	10	√
13	El Emam, Khaled	•	•	•	•	2	1	1	2	14	•	•	•	147
14	Fenton, Norman	10	13	•	13	•	•	•	•	•	•	•	•	√
15	Fenwick, Peter	•	•	•	•	•	14	•	•	•	•	•	•	•
16	Fuggetta Alfonso	9	13	13	11	•	•	•	•	•	•	•	•	343

Table 7 (continued)

No.	Name	Rankings												
		(Glass 1998)	(Glass 1999)	(Glass 2000)	(Glass and Chen 2001)	(Glass and Chen 2002)	(Glass and Chen 2003)	(Glass and Chen 2005)	(Tse et al. 2006)	(Eric Wong et al. 2008)	(Eric Wong et al. 2009)	(Wong et al. 2011)	(Wong et al. 2011)	Ioannidis et al.
17	Glass, Robert L.	5	3	5	5	5	4	4	10	14	•	•	•	•
18	Harman, Mark	•	•	•	•	•	•	•	•	14	•	9	11	8
19	Harrold, Mary Jean	14	•	•	•	14	•	9	•	•	•	•	•	20
20	Hatton, Les	14	15	•	•	•	•	•	•	•	•	•	•	•
21	He, Xudong	•	•	•	•	•	•	13	•	•	•	•	•	•
22	Henderson-Sellers, Brian	•	•	•	•	12	11	6	•	•	•	•	•	137
23	Huang, Chin-Yu	•	•	•	•	•	•	•	•	•	14	8	7	226
24	Hutton, Les	•	•	•	•	14	•	•	•	•	•	•	•	•
25	Jackson, Michael	•	6	6	10	12	•	•	•	•	•	•	•	149
26	Jorgensen, Magne	•	•	•	•	•	•	•	3	1	1	1	1	62
27	Kim, Ho Myoung	•	•	•	•	•	•	11	7	9	13	•	•	•
28	Kim, Hyoung-Joo	•	•	•	•	•	5	3	8	10	•	•	•	•
29	Kitchenham, Barbara	•	15	9	7	3	3	2	4	2	5	5	5	15
30	Lai, Richard	2	1	1	1	1	2	7	•	•	10	15	•	•
31	Ling, Tok Wang	12	•	•	•	•	•	•	•	•	•	•	•	•

Table 7 (continued)

No.	Name	Rankings												
		(Glass 1998)	(Glass 1999)	(Glass 2000)	(Glass and Chen 2001)	(Glass and Chen 2002)	(Glass and Chen 2003)	(Glass and Chen 2005)	(Tse et al. 2006)	(Eric Wong et al. 2008)	(Eric Wong et al. 2009)	(Wong et al. 2011)	(Wong et al. 2011)	Ioannidis et al.
32	Liutz, Robyn R.	•	•	•	•	•	•	•	•	•	•	7	•	432
33	Memon, Atif M.	•	•	•	•	•	•	•	•	•	•	•	8	106
34	Miller, James	•	•	•	•	•	•	•	8	10	9	11	12	318
35	Nesi, Paolo	•	•	•	•	11	9	•	•	•	•	•	•	•
36	Pfleeger, Shari Lawrence	6	10	•	•	•	•	•	•	•	•	•	•	167
37	Piattini, Mario	•	•	•	•	•	•	•	•	•	•	•	13	•
38	Poore, J. H.	•	•	•	15	14	14	•	13	•	•	•	•	•
39	Rueson, Per	•	•	•	•	•	•	13	12	•	5	4	•	122
40	Saiedian, Hossein	10	•	•	•	•	•	•	•	•	•	•	•	•
41	Saleh, Kassem	8	9	3	3	4	5	8	•	•	•	•	•	•
42	Shin, Kang G.	12	•	•	•	•	•	•	•	•	•	•	•	√
43	Sjøberg, Dag I. K.	•	•	•	•	•	•	•	•	•	•	11	•	276
44	Stamelos, Ioannis	•	•	•	•	•	•	•	•	12	11	•	•	•
45	Tian, Jeff	14	11	8	•	•	•	•	11	7	12	14	•	•
46	Tiwana, Amrit	•	•	•	•	•	•	•	•	•	•	•	13	√
47	Visaggio, Giuseppe	•	•	15	•	•	•	•	•	•	•	•	•	•
48	von Vliet, Hans	•	•	•	•	•	•	•	•	•	•	•	13	•

Table 7 (continued)

No.	Name	Rankings												
		(Glass 1998)	(Glass 1999)	(Glass 2000)	(Glass and Chen 2001)	(Glass and Chen 2002)	(Glass and Chen 2003)	(Glass and Chen 2005)	(Tse et al. 2006)	(Eric Wong et al. 2008)	(Eric Wong et al. 2009)	(Wong et al. 2011)	(Wong et al. 2011)	Ioannidis et al.
49	Weyuker, Elaine J.	3	5	10	6	•	•	•	•	•	•	•	•	36
50	Whittaker, James	•	•	•	•	•	12	13	13	•	•	•	•	•
51	Wohlin, Claes	•	•	•	•	•	14	10	•	•	8	9	•	53
52	Wong, Johnny S. K.	1	2	2	4	5	•	•	•	•	•	•	•	•
53	Wu, Pei-Chi	•	•	•	•	•	14	•	•	•	•	•	•	•
54	Zhuge, Hai	•	•	•	•	•	•	13	1	2	2	3	3	√

Next, we compare the most recent JSS (Karanatsiou et al. 2019) ranking with Ioannidis et al.'s database. The newest JSS ranking is analyzed separately as the methodology for ranking the scientists has changed. Key changes are:

- The authors did not focus on the predefined set of journals from the earlier rankings but substantially extended the set of journals considered. They also included conferences.
- Researcher productivity was now measured as total paper count.
- Researchers were assigned to three groups based on their career age (experienced researchers with more than 12 years of research, consolidated researchers with 8–12 years of research, and early-stage researchers with up to 7 years of research)
- New analyses were included by considering several types of rankings based on (a) top journal publications, (b) top publications (including journals and conferences), and (c) impact measured in citations.

For the comparison with the database, we only considered the list of experienced researchers, given that the database covered the period of 1960–2019 and mainly included researchers in the experienced category (see Fig. 4). Overall, 45 researchers classified as experienced were listed in the last JSS ranking. Of these 45 researchers, 28 (62%) were included in the Ioannidis et al. database, all classified correctly as software engineering researchers.

Ranking by Fernandes (2014)*:* The ranking by Fernandes assessed researchers by productivity. The data was obtained by focusing on the DBLP database (Ley 2009). Fernandes identified 30 top researchers in the period 1971–2012. Given that a more extended period is covered, this ranking is a particularly interesting reference for comparison. Of the 30 researchers, 24 were included in the database and classified as software engineering researchers. One researcher ranked by Fernandes (Taghi M. Khoshgoftaar) was found in the Ioannidis et al. database but not classified as software engineering in his main field. Overall, 83% of the researchers in the ranking by Fernandes are also included in the database.

Fernandes used fractional and harmonic citations (Hagen 2014), which consider relative author contributions. However, no impact (citations) was measured. In the Composite Citation index used by Ioannidis, also author contributions are distinguished. It is noteworthy that one measure (*ncsf:* total citations to papers for which the researcher is the sole or the first author) used would not apply to software engineering. In software engineering, the last author position would usually be considered the least contributing, while in the medical field, the last author position stands out (Baerlocher et al. 2007).

Overall, the analysis of the database for software engineering shows that when using the database in another field, it may be necessary to investigate researchers' classification. Researchers wrongly classified as being in the investigated field may be easier to detect than researchers not appearing in the field due to wrong classifications. We used existing rankings to detect researchers not appearing in software engineering. However, previous rankings may not be available in other fields, which would make the detection of missing researchers more difficult.

Conclusion

We analyzed the top-ranked software engineering researchers in the database published by Ioannidis et al. The database provided the opportunity to gain novel insights on the top authors in software engineering, as a new measure was used to identify the authors (Composite Citation index), and that the ranking was considering the number of self-citations. Furthermore, the database included other fields that were assessed using the same measures. Therefore, software engineering researchers could learn how they compare to other fields.

We formulated three research questions, which are briefly answered as follows:

RQ1: How do author citations in software engineering compare to other fields concerning the citation measures provided by Ioannidis et al.? We first compared software engineering with nine other related fields in the area of Information & Communication Technology, where it was ranked third concerning Compound Citations. After that, we compared software engineering with 20 other fields included in the database. Software engineering lies in the center of the fields for Total Papers, Total Cites, H-Index, and Composite Citation index. Given that 37% of the researchers were wrongly classified as software engineers in their main field, there is an effect on the comparison. The effect was minor for most measures, except the percentage of self-citations.

What are the top authors' demographics (e.g., country of origin, institutions, etc.)? In total, Ioannidis et al. identified 441 researchers as top researchers in software engineering. However, after examining the authors' list in further detail, only 278 authors were in software engineering. By looking at other rankings done in software engineering, we identified seven additional top researchers that were not classified as software engineers in their main field. The top researcher in the field of software engineering is Barry Boehm. In our reporting, we also highlighted those researchers in the ranking that are software engineers. The top countries and institutions were also identified. The top countries were the USA, Canada, and the UK. The institution with most software engineering top authors was Microsoft. We also found that career age and productivity were not associated with the Composite Citation index, which may be also attributed to not accounting for publication age, which was not included in the ranking.

In which other fields software engineering authors are active? In total, 33 fields were listed as second fields for researchers classified as software engineers in their main field. The most frequent field that occurred together with software engineering was Artificial Intelligence & Image Processing, followed by Human Factors.

Overall, given that a large portion of researchers classified as software engineers were not in software engineering, we suggest that other fields using the database for their rankings need to review the author lists. As the classification was done using convolutional neural networks, there may always be a portion of not correctly classified researchers. However, overall, 60%–80% of the researchers listed in other software engineering rankings were identified, there is some confidence in the identification of relevant researchers.

In future work, we suggest using the database further to assess researchers in other fields, e.g., taking into account the conventions in the field regarding the author order as weighted in the Composite Citation index and the reliability of underlying classification of the researchers.

Table 8 Top author list - software engineering

Rank	Author
1	Terzopoulos, Demetri
2	Boehm, Barry*
3	Levoy, Marc
4	Hoppe, Hugues
5	Guibas, Leonidas
6	Basili, Victor*
7	Slater, Mel
8	Harman, Mark*
9	Keim, Daniel
10	Parnas, David Lorge*
11	Briand, Lionel*
12	Wadler, Philip*
13	Witkin, Andrew
14	Kobbelt, Leif
15	Kitchenham, Barbara*
16	Turk, Greg
17	Godefroid, Patrice*
18	Hertzmann, Aaron
19	Fedkiw, Ronald
20	Harrold, Mary Jean*
21	Taubin, Gabriel
22	Hanrahan, Pat
23	Funkhouser, Thomas
24	Schmidt, Douglas C.*
25	Ware, Colin
26	Gleicher, Michael*
27	Nielsen, Jakob*
28	Zeller, Andreas*
29	Lam, Monica S.*
30	Rusinkiewicz, Szymon
31	Herbsleb, Jim*
32	Debevec, Paul
33	Van Lamsweerde, Axel*
34	Holzmann, Gerard J.*
35	Farouki, Rida T.
36	Weyuker, Elaine J.*
37	Manocha, Dinesh
38	Durand, Frédo
39	Ball, Thomas*
40	Van Wijk, Jarke
41	Hassan, Ahmed E.*
42	Garlan, David*
43	Seidel, Hans Peter
44	Brooks, Frederick P.*
45	Cohen-Or, Daniel
46	Sen, Koushik*

Table 8 (continued)

Rank	Author
47	Raskar, Ramesh
48	Alexa, Marc
49	Taylor, Richard*
50	Shewchuk, Jonathan Richard
51	Fattal, R.
52	Cohen, Michael F.
53	Wohlin, Claes*
54	Stasko, John
55	Bosch, Jan*
56	Gross, Markus
57	Hodgins, Jessica
58	Flanagan, Cormac*
59	Ma, Kwan Liu*
60	Kemerer, Chris F.*
61	Fitzgerald, Brian*
62	Jørgensen, Magne*
63	Sederberg, Thomas W.
64	Batory, Don*
65	Pugh, William*
66	Azuma, Ron
67	Reynolds, John C.*
68	Ramamoorthi, Ravi
69	Necula, George*
70	Meyer, Bertrand*
71	Sutcliffe, Alistair*
72	Nuseibeh, Bashar*
73	Pottmann, Helmut
74	Thomas, D. A.
75	Dybå, Tore*
76	Leroy, Xavier*
77	Larus, James*
78	Tonella, Paolo*
79	Aiken, Alex*
80	Rinard, Martin*
81	Stam, Jos
82	Bajaj, Chandrajit
83	Kruchten, Philippe*
84	Levin, David
85	Kramer, Jeffrey*
86	Zimmermann, Thomas*
87	Mockus, Audris*
88	Myers, Andrew C.*
89	Garland, Michael*
90	Williams, Laurie*
91	Mens, Tom*

Table 8 (continued)

Rank	Author
92	Gulwani, Sumit*
93	Heer, Jeffrey*
94	Wilkinson, Leland
95	Jackson, Daniel*
96	Korel, Bogdan*
97	Shamir, Ariel
98	Rossignac, Jarek
99	Shaw, Mary*
100	Holt, R. C.*
101	Appel, Andrew W.*
102	Perlin, Ken*
103	Shepperd, Martin*
104	Spinellis, Diomidis*
105	Kass, Michael
106	Memon, Atif*
107	Barr, Alan H.
108	Blinn, James F.
109	Robillard, Martin P.*
110	Czarnecki, Krzysztof*
111	Gupta, Rajiv*
112	De Lucia, Andrea*
113	Curless, Brian*
114	Kaufman, Arie
115	Johnson, Ralph
116	Storey, Margaret Anne*
117	Binkley, David
118	Lehman, M. M.*
119	Medvidovic, Nenad*
120	Owens, John D.
121	Menzies, Tim*
122	Runeson, Per*
123	Arcuri, Andrea*
124	Devanbu, Premkumar*
125	Lévy, Bruno
126	Lanza, Michele*
127	Xie, Tao*
128	Sheffer, Alla
129	Wattenberg, Martin*
130	Ward, Gregory J.
131	Seli_, Bran*
132	Murphy, Gail C.*
133	Munzner, Tamara
134	Thalmann, Daniel
135	Kiczales, Gregor*
136	Felleisen, Matthias*

Table 8 (continued)

Rank	Author
137	Henderson-Sellers, Brian*
138	Pai, Dinesh K.
139	Aho, Alfred*
140	Igarashi, Takeo
141	Egyed, Alexander
142	Purchase, Helen
143	Fekete, Jean Daniel*
144	Baraff, David
145	Zorin, Denis
146	McMinn, Phil*
147	El Emam, Khaled*
148	Pfister, Hanspeter
149	Tsai, Wei Tek*
150	Hu, Shi Min
151	Reiss, Steven*
152	Carpendale, Sheelagh*
153	Heckbert, Paul S.*
154	Hierons, Robert M.*
155	Schröder, Peter
156	Canfora, Gerardo*
157	Horwitz, Susan*
158	Gotsman, Craig J.
159	Farin, Gerald
160	Müller, Matthias*
161	Harris, Tim*
162	Ernst, Michael D.*
163	Elber, Gershon
164	Lindstrom, Peter
165	Eppler, Martin J.
166	Nagappan, Nachiappan*
167	Pfleeger, Shari Lawrence*
168	Easterbrook, Steve*
169	Offutt, A. J.
170	Jensen, Henrik Wann
171	Offutt, Jeff*
172	Agarwala, Aseem
173	Chambers, Craig
174	Amenta, Nina
175	Weimer, Westley*
176	Sorkine, Olga*
177	Dodgson, Neil A.
178	Luebke, David
179	Sommerville, Ian*
180	Desbrun, Mathieu
181	Finkelstein, Anthony*

Table 8 (continued)

Rank	Author
182	Yu, Yizhou
183	Ebert, Christof*
184	Westermann, Rüdiger
185	Cleland-Huang, Jane*
186	Littlewood, Bev*
187	Koschke, Rainer*
188	Hughes, John*
189	Erwig, Martin*
190	Popovi_, Zoran
191	Paris, Sylvain
192	Antoniol, Giuliano*
193	Johnson, Chris*
194	Kajiya, Jim
195	North, Chris*
196	Pinkall, Ulrich
197	Tichy, Walter F.*
198	Ryder, Barbara G.*
199	McKinley, Kathryn S.*
200	Ju, Tao
201	Anton, Annie*
202	Wong, Kenny*
203	Visser, Eelco*
204	Abrahamsson, Pekka*
205	Warren, Joe
206	Hart, J.
207	Kuhn, D. Richard*
208	Bertolino, Antonia*
209	Zyda, Michael
210	Maiden, Neil*
211	Krinke, Jens*
212	Bridson, Robert
213	Badler, Norman I.
214	Rosenblum, David S.*
215	Scacchi, Walt*
216	Seaman, Carolyn*
217	Curtis, Bill*
218	van Deursen, Arie*
219	Lipman, Yaron
220	Schmalstieg, Dieter
221	Zave, Pamela*
222	Nielson, Gregory M.
223	Max, Nelson
224	Pajarola, Renato
225	Foxlin, Eric
226	Huang, Chin Yu*

Table 8 (continued)

Rank	Author
227	Vouk, Mladen Alan*
228	Jüttler, Bert
229	Fuchs, Henry
230	Odersky, Martin*
231	Orso, Alessandro*
232	Palsberg, Jens*
233	Chen, Tsong Yueh*
234	Magee, Jeff*
235	Wei, Li Yi*
236	Greenberg, Donald P.
237	Walker, David*
238	Prechelt, Lutz*
239	Cordy, James R.*
240	Harper, Robert*
241	Tai, Kuo Chung*
242	Rieser, John J.
243	Franke, R.
244	Jones, Simon Peyton*
245	Hodges, Larry F.*
246	Hölzle, Urs*
247	LaViola, Joseph J.*
248	Steed, Anthony
249	Jackson, Michael*
250	Hamann, Bernd
251	Magnenat-Thalmann, Nadia*
252	Lécuyer, Anatole*
253	Wand, Mitchell*
254	France, Robert B.*
255	Rothermel, Gregg*
256	Griswold, William G.*
257	Eick, Stephen G.*
258	Mei, Hong*
259	Whittle, Jon*
260	Kazman, Rick*
261	Marcus, Andrian*
262	Ko, Andrew J.
263	Petersen, Kai*
264	Ungar, David*
265	Sneed, Harry M.*
266	Pascucci, Valerio
267	Knight, John*
268	Rajlich, Václav*
269	Perry, Dewayne E.*
270	Dobkin, David P.
271	Di Penta, Massimiliano*

Table 8 (continued)

Rank	Author
272	Lewis, J. P.
273	Poshyvanyk, Denys*
274	Murugesan, San*
275	Ducasse, Stéphane*
276	Sjoberg, Dag I.K.*
277	Lee, Seungyong
278	Weiskopf, Daniel
279	DeRose, Tony
280	Mendes, Emilia*
281	Lämmel, Ralf*
282	Schneidewind, N. F.*
283	Grundy, John*
284	Dingsøyr, Torgeir*
285	Fuggetta, Alfonso*
286	Chang, Shi Kuo*
287	Nierstrasz, Oscar*
288	Leung, Hareton*
289	Lethbridge, Timothy C.*
290	Shirley, Peter
291	Wong, W. Eric*
292	Boehm, Hans J.*
293	Wald, Ingo
294	Sewell, Peter*
295	Osterweil, Leon J.*
296	Corbett, James C.*
297	Clarke, Lori A.*
298	Snyder, John
299	Ellis, Stephen R.
300	Peters, Jörg
301	Adve, Vikram*
302	Bird, Christian*
303	Apel, Sven*
304	Brun, Yuriy*
305	Nishita, Tomoyuki
306	German, Daniel M.*
307	Woodside, Murray*
308	Goldman, Ron*
309	Grossman, Dan*
310	Fisher, Danyel
311	Reinhard, Erik
312	Bieman, James M.*
313	Guo, Philip J.
314	Peyton Jones, S. I.M.O.N.*
315	Pauly, Mark
316	Ramalingam, G.*

Table 8 (continued)

Rank	Author
317	Potts, Colin*
318	Miller, James
319	Damian, Daniela*
320	Laprie, J. C.*
321	Ruhe, Guenther*
322	Hughes, John F.*
323	Pierce, Benjamin C.*
324	Cooper, Keith*
325	Ashikhmin, Michael
326	Petrank, Erez*
327	Debray, Saumya*
328	Burnett, Margaret*
329	Cugola, Gianpaolo
330	Deussen, Oliver
331	Van de Panne, Michiel
332	Babar, Muhammad Ali*
333	Shen, Han Wei
334	Emmerich, Wolfgang*
335	Ghezzi, Carlo*
336	Martin, Ralph R.
337	Ammann, Paul*
338	Gomaa, Hassan*
339	Visser, Willem*
340	Wirth, N.*
341	Qadeer, Shaz*
342	Matusik, Wojciech
343	Agrawala, Maneesh
344	McCabe, Thomas J.*
345	Bodden, Eric*
346	Catmull, Edwin
347	O'Brien, James F.
348	Lin, Ming
349	Shull, Forrest*
350	Ribarsky, William
351	Fraser, Gordon*
352	Kacker, Raghu N.*
353	Bacon, David F.*
354	DeLine, Robert*
355	Zhu, Hong
356	Guo, Baining
357	Cook, William R.*
358	Chilimbi, Trishul
359	Laidlaw, David H.
360	Zhang, Hongyu
361	Mezini, Mira*

Table 8 (continued)

Rank	Author
362	Kazhdan, Misha
363	Thomas, Bruce H.
364	Niazi, Mahmood*
365	Kautz, Jan
366	Regehr, John*
367	Zhou, Kun
368	Benton, Nick*
369	Singh, Karan
370	Guéhéneuc, Yann Gaël*
371	Amarasinghe, Saman*
372	Hauser, Helwig
373	Davis, Alan M.*
374	Hudak, Paul*
375	Magnor, Marcus
376	Cignoni, Paolo
377	Aldrich, Jonathan*
378	Andrienko, Gennady
379	Sheard, Tim*
380	Malaiya, Yashwant K.*
381	Yau, Stephen S.*
382	Bultan, Tevfik*
383	Lavoué, Guillaume
384	Loop, Charles
385	Gokhale, Swapna*
386	Gorton, Ian*
387	Dwyer, Matthew B.*
388	Musa, John D.*
389	Grunske, Lars*
390	Mitra, Niloy J.
391	Ostromoukhov, V.
392	Firesmith, Donald G.*
393	Interrante, Victoria
394	Marchesi, M.*
395	Rountev, Atanas
396	Aliaga, Daniel G.
397	Diehl, Stephan*
398	Abran, Alain*
399	Moss, J. Eliot B.*
400	Parker, Steven
401	Notkin, David*
402	Lee, Jehee
403	Lischinski, Dani
404	Seidman, Stephen*
405	Smaragdakis, Yannis*
406	Van Der Hoek, Andre*

Table 8 (continued)

Rank	Author
407	Miller, Gavin S.P.
408	Glinz, Martin*
409	Frankl, Phyllis*
410	Dykes, Jason
411	Hinze, Ralf*
412	Meijer, Erik*
413	Lopes, Cristina V.*
414	Hindle, Abram*
415	Decarlo, Doug
416	Ebert, David S.
417	Goel, Amrit L.*
418	Heitmeyer, Constance*
419	Sillion, Francois
420	Noble, James*
421	Franz, Michael*
422	Xi, Hongwei*
423	Yahav, Eran*
424	Ferwerda, James
425	Yoo, Shin*
426	Consel, Charles*
427	De Floriani, Leila
428	Jézéquel, Jean Marc*
429	Wagner, Stefan*
430	Lo, David*
431	Franch, Xavier*
432	Lutz, Robyn*
433	Schmid, Klaus*
434	Telea, Alexandru C.*
435	Biffl, Stefan*
436	Patané, Giuseppe
437	Kunii, Tosiyasu L.*
438	Garousi, Vahid*
439	Zdun, Uwe*
440	Kim, Kane*
441	Roy, Chanchal K.*

Complete list of top-listed software engineering researchers

The following table provides the list of top researchers identified for software engineering.

Funding Open access funding provided by Blekinge Institute of Technology.

References

Baerlocher, M. O., Newton, M., Gautam, T., Tomlinson, G., & Detsky, A. S. (2007). The meaning of author order in medical research. *Journal of Investigative Medicine, 55*(4), 174–180.

Bin Ali N., Edison H., & Torkar R. (2020). The impact of a proposal for innovation measurement in the software industry. In: Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), pp. 1–6, Bari Italy, ACM.

Bornmann, L., & Daniel, H. D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation, 64*(1), 45–80.

Croux, Christophe, & Dehon, Catherine. (2010). Influence functions of the spearman and kendall correlation measures. *Statistical Methods & Applications, 19*(4), 497–515.

Fernandes, João. M. (2014). Authorship trends in software engineering. *Scientometrics, 101*(1), 257–271.

Fernandes, João. M., & Monteiro, Miguel P. (2017). Evolution in the number of authors of computer science publications. *Scientometrics, 110*(2), 529–539.

Garousi V., & Fernandes J.M. (2016). Highly-cited papers in software engineering: The top-100. p. 21: Information and Software Technology.

Garousi, V., & Fernandes, J. M. (2017). Quantity versus impact of software engineering papers: A quantitative study. *Scientometrics, 112*(2), 963–1006.

Glass, R. L. (1994). An assessment of systems and software engineering scholars and institutions. *The Journal of Systems and Software, 27*(1), 63–67.

Glass, R. L. (1995). Editor's corner an assessment of systems and software engineering scholars and institutions, 1993 and 1994. *Journal of System and Software, 31*(1), 3–6.

Glass, Robert L. (1996). An assessment of systems and software engineering scholars and institutions (1993–1995). *Journal of System and Software, 35*(1), 85–89.

Glass, R. L. (1998). An assessment of systems and software engineering scholars and institutions (1993–1997). *Journal of System and Software, 43*(1), 59–64.

Glass, R. L. (1999). An assessment of systems and software engineering scholars and institutions (1994–1998). *Journal of System and Software, 49*(1), 81–86.

Glass, Robert L. (2000). An assessment of systems and software engineering scholars and institutions (1995–1999). *Journal of System and Software, 54*(1), 77–82.

Glass, Robert L., & Chen, Tsong Yueh. (2001). An assessment of systems and software engineering scholars and institutions (1996–2000). *Journal of System and Software, 59*(1), 107–113.

Glass, Robert L., & Chen, Tsong Yueh. (2002). An assessment of systems and software engineering scholars and institutions (1997–2001). *Journal of System and Software, 64*(1), 79–86.

Glass, R. L., & Chen, T. Y. (2003). An assessment of systems and software engineering scholars and institutions (1998–2002). *Journal of System and Software, 68*(1), 77–84.

Glass, Robert L., & Chen, Tsong Yueh. (2005). An assessment of systems and software engineering scholars and institutions (1999–2003). *Journal of System and Software, 76*(1), 91–97.

Hagen, Nils T. (2014). Counting and comparing publication output with and without equalizing and inflationary bias. *Journal of Informetrics, 8*(2), 310–317.

Ioannidis, J. P. A., Boyack, K. W., & Baas, J. (2020). Updated science-wide author databases of standardized citation indicators. *PLOS Biology, 18(10), 1–3*, 10.

Ioannidis, John P.A.., Klavans, Richard, & Boyack, Kevin W. (2016). Multiple citation indicators and their composite across scientific disciplines. *PLOS Biology, 14*(7), e1002501.

Karanatsiou, D., Li, Y., Arvanitou, E. M., Misirlis, N., & Wong, W. E. (2019). A bibliometric assessment of software engineering scholars and institutions (2010–2017). *Journal of Systems and Software, 147*, 246–261.

Ley, Michael. (2009). Dblp: Some lessons learned. *Proceedings of the VLDB Endowment, 2*(2), 1493–1500.

Mathew, G., Agrawal, A., & Menzies, T. (2019). Finding trends in software research. *IEEE Transactions on Software Engineering*. https://doi.org/10.1109/TSE.2018.2870388
Molléri, J. S., Petersen, K., & Mendes, E. (2018). Towards understanding the relation between citations and research quality in software engineering studies. *Scientometrics, 117*(3), 1453–1478.
Pendlebury, D. A., & Adams, J. (2012). Comments on a critique of the thomson reuters journal impact factor. *Scientometrics, 92*(2), 395–401.
Poulding, S., Petersen K., Feldt R., & Garousi, V. (2015). In: Using citation behavior to rethink academic impact in software engineering.
Schreiber, Michael. (2008). A modification of the h-index: The hm-index accounts for multi-authored manuscripts. *Journal of Informetrics, 2*(3), 211–216.
Teufel, S., Siddharthan, A., & Tidhar, D. (2006). An annotation scheme for citation function. Sydney, Australia: Association for Computational Linguistics. pp. 80.
Tse, T. H., Chen, T. Y., & Glass, R. L. (2006). An assessment of systems and software engineering scholars and institutions (2000–2004). *Journal of System and Software, 79*(6), 816–819.
Wong, W. E., Tse, T. H., Glass, R. L., Basili, V. R., & Chen, T. Y. (2008). An assessment of systems and software engineering scholars and institutions (2001–2005). *Journal of System and Software, 81*(6), 1059–1062.
Wong, W. E., Tse, T. H., Glass, R. L., Basili, V. R., & Chen, T. Y. (2009). An assessment of systems and software engineering scholars and institutions (2002–2006). *Journal of System and Software, 82*(8), 1370–1373.
Wong, W. E., Tse, T. H., Glass, R. L., Basili, V. R., & Chen, T. Y. (2011). An assessment of systems and software engineering scholars and institutions (2003–2007 and 2004–2008). *Journal of Systems and Software, 84*(1), 162–168.