DOI:10.1145/1941487.1941517

Most of a computer scientist's production can go uncounted if a standard bibliographic service is used.

BY JACQUES WAINER, SIOME GOLDENSTEIN, AND CLEO BILLA

Invisible Work in Standard Bibliometric Evaluation of Computer Science

MULTIDISCIPLINARY COMMITTEES ROUTINELY make strategic decisions, rule on subjects ranging from faculty promotion to grant awards, and rank and compare scientists. Though they may use different criteria for evaluations in subjects as disparate as history and medicine, it seems logical for academic institutions to group together mathematics, computer science, and electrical engineering for comparative evaluation by these committees.

These evaluations will be more frequent as the number of scientists increases. Since funding sources grow more slowly, and research practices vary among different subjects, using the same criteria in different areas may produce notable injustices. The ongoing discussion on CS research evaluation[4,6] builds the case for the CS community defending itself from expected injustices in its future comparative evaluations.

The traditional assessment criteria are based on Thomson Reuters' Web of Science (WoS) indexing service, quantifying the production and number of citations of individual scientists, university departments (such as was covered in Torres-Salinas et al.[5]), whole universities,[2] countries,[7] and scientific areas.

Computer scientists have an intuitive understanding that these assessment criteria are unfair to CS as a whole. Here, we provide some quantitative evidence of such unfairness.

We define researchers' invisible work as an estimation of all their scientific publications not indexed by WoS or Scopus. Thus, the work is not counted as part of scientists' standard bibliometric evaluations. To compare CS invisible work to that of physics, mathematics, and electrical engineering, we generated a controlled sample of 50 scientists from each of these fields from top U.S. universities and focused on the distribution of invisible work rate for each of them using statistical tests.

We defined invisible work as the difference between number of publications scientists themselves list on their personal Web pages and/or publicly available curriculum vitae (we call their "listed production") and number of publications listed for the same scientists in WoS and Scopus. The invisible

» key insights

- **On average, 66% of a computer scientist's published work is not accounted for in WoS.**
- **This rate is much higher than for physics and math researchers.**
- **Researchers in the various subareas within CS have different rates, relatively higher for those in AI, architecture, and systems.**

work rate is the invisible work divided by number of listed production. Note that our evaluation of invisible work rate is an approximation of the true invisible work rate because the listed production of particular scientists may not include all of their publications.

WoS and Scopus

Thomson Reuters WoS[a] is the most traditional citation-indexing service and standard in most bibliometric research. In 2008, WoS indexed 393 journals in CS. The CS community is well aware of WoS limitations:

Incomplete coverage. Before 2008, WoS did not include conferences other than those published in Springer's *Lecture Notes in Computer Science* series. Even though in mid-2009 Thompson Reuters introduced a new service that indexes around 12,000 conference proceedings, it is unclear today if most science-evaluation bodies will use (and pay for) this new service;

Missing important titles. WoS does not include journals considered important in the field, including *Transactions on Algorithms* and *Journal of Experimental Algorithms*, both from ACM; and

Not all CS. WoS includes within CS some journals most of the CS community would not consider CS journals, including the *Journal of Chemical Information and Modeling*, which published 222 articles in 2008 compared to the *Journal of the ACM*, which published 28 articles in 2008.[b]

a http://www.isiknowledge.com
b Data from *Journal of Citation Reports 2008*; http://admin-apps.isiknowledge.com/JCR/JCR?PointOfEntry=Home

For this article, we collected data for both the standard WoS indexing (here called "WoS") and the new WoS with Proceedings (here called "WoS-P"). Elsevier developed Scopus[c], a recent competitor of WoS, including 849 CS journals and 162 conferences in 2009[d]. Some articles compare the differences in coverage and citation counts between WoS and Scopus.[1,3] In particular, Bar-Ilan[1] indicated that for CS, the h-index calculated using Scopus is on average 13% higher than the one calculated using WoS, based on a sample of 12 highly cited Israeli CS researchers.

Methodology

Selection of scientists and their listed production. Based on the 2009 ranking of the best graduate schools in the U.S. as determined by *U.S. News & World Report*[e], we selected the universities ranked sixth to 10th in physics, math, CS, and electrical, electronic, and communications engineering. The rationale for skipping the top five is they could have specific features in their research culture that would place them as outliers within their field. Well-ranked departments probably reflect a balanced distribution of researchers in the subareas in physics, math, EE, and CS and are, hopefully, representative of the balance of these subareas in the general research population. We randomly selected 10 faculty members from each department with up-to-date "Publications" sections on their Web pages, totaling 200 researchers, 50 from each area. For these researchers, we counted the number of publications (excluding technical reports) from 2000 to 2007, inclusive, considering this number to be a researcher's listed production. Of the CS researchers, 10 divided their production into journal papers, conference papers, and other forms of production, for which we recorded the partial totals as well.

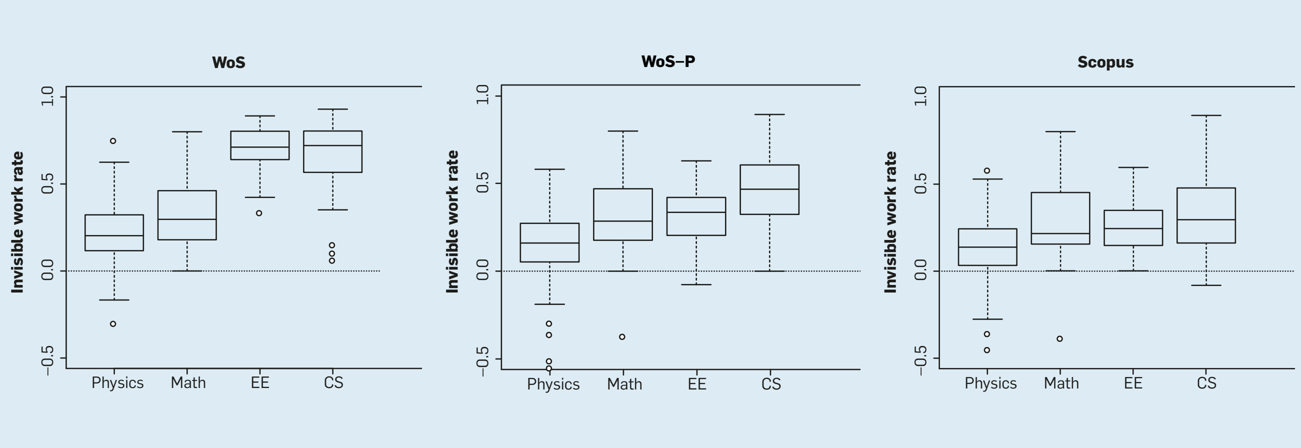
c http://www.scopus.com
d http://www.scopus.com/scopus/source/browse.url?subjectArea=1700
e http://www.usnews.com/sections/rankings

We collected the WoS and Scopus data in October 2009 and the WoS-P in February 2010.

Indexing services data. To collect the WoS data, we used the `Author finder` option in the basic search; for the Scopus data we used the `Author search` option. Since data collection involved subjective decisions, particularly relating to homonyms, at least two authors of this article independently collected the data for each researcher and discussed these subjective decisions until there were no discrepancies. For each service, we collected the number of journal and conference papers for each researcher, considering this sum to be the researcher's total production. Scopus alternates its classification of LNCS papers, sometimes labeling them as a conference and sometimes as a journal publication. To identify and count the number of journal articles from a particular CS researcher, we excluded all papers published in LNCS. The number of conference papers is total number of publications minus number of journal articles.

Metric. We defined the invisible work rate for a researcher as

Boxplots of the invisible work rate for each scientific area based on WoS, WoS-P, and Scopus data.



invisible work rate =
(number of listed papers − number of papers in indexing service)
÷ number of listed papers

and analyzed the distribution of invisible work rates for the various scientific areas. The number of listed papers is the number of publications researchers publicly listed as their scientific production. An invisible work rate of 1 means none of a scientist's publications are accounted for in the indexing service. A rate of 0 is generally understood to indicate that all of a scientist's publications are represented in the indexing service, but the practical meaning is less clear; the number of publications listed by a researcher is the same as the number of entries in the bibliometric service. Some works on a researcher's list of publications may not be represented in the service, but the number is balanced by the same number of references in the service but not in the researcher's list of publications.

There is no consensus among researchers regarding which publications to list on one's Web pages; for example, some CS researchers do not include editorials, invited papers, or other non-peer-reviewed material. However, an indexing service may not make such a distinction. Similarly, researchers in other areas may not list conference papers, but some conferences are indexed by WoS-P and Scopus. Researchers may thus have papers attributed to them by the service that are not listed on their publication pages. Also, some researchers, especially the more-prolific ones, may truly forget to add one or another paper to their lists of publications. We also discovered that one of these services had made at least one error—duplication of the same paper title attributed to two different journals, with only one correct.

In light of these facts, the bibliometric services count some papers that do not appear on a researcher's publication list, potentially leading to a negative invisible work rate. We kept the negative rates in our research since our goal was to estimate the number of published CS papers left out of the standard indexing services, and the difference between what each area considers "published work" is part of the variability across different areas.

Statistical analysis. We say "significant" to indicate the differences are statistically significant with 95% confidence. If not indicated, we determined significance was determined by non-paired t-test and report only the p-value of the test. In a few cases, where indicated, we used a paired t-test to compare the change in the average CS invisible work rate between two different indexing services and a one-way ANOVA test to compare the invisible work rate of assistant, associate, and full professors.

Results

The boxplots in the figure here outline the distribution of the invisible work rate for each area using WoS, WoS-P, and Scopus data, respectively; a boxplot shows the lower quartile (the value cutting off the lowest 25% of data), the median, and the upper quartile of a distribution (the box part of the plot), and the maximum and minimum (the ⊤ and ⊥ lines). Outliers in a boxplot, shown as circles, are defined as data lying more than 1.5 IQR below the lower quartile or 1.5 IQR above the upper quartile, where IQR is the interquartile range, or distance between the lower and the upper quartiles. Table 1 lists the data in the figure as the mean (with standard deviation in parenthesis) of the invisible work for each area and service.

For WoS, as expected, fewer of the physics and math researchers' papers were unaccounted for. The average invisible work rate for CS is significantly higher than for physics (p-value < 2.2e-16) and math (p-value = 8.1e-12). EE had a higher invisible rate than CS, but the difference was not significant (p-value = 0.1725). For WoS-P, the CS average invisible work rate was significantly higher than physics, math, and EE (p-value = 1.163e-09, p-value = 0.001776, and p-value = 0.0005013, respectively). For Scopus, CS was significantly higher than physics and EE (p-value = 1.789e-05 and p-value = 0.02799, respectively) but not than math (p-value = 0.38).

For all four areas, WoS-P was better than WoS, and Scopus was better than WoS-P in representing the work of researchers. Within CS, the invisible rate significantly improved when we used WoS-P instead of WoS (paired t-test p-value = 2.229e-14). It also improved significantly when we used Scopus instead of WoS-P (paired t-test p-value = 0.0001222).

One expected reason for the large invisible rate in CS was the emphasis CS puts on publishing in conference proceedings. We evaluate the invisible work rate for conferences and journals separately in the following sections, exploring whether researchers in the subareas within CS and at different moments in their careers have different invisible work rates.

Conferences vs. journals. Following the 10 CS researchers who partitioned their publications between conferences and journals, we calculated the invisible work rate for conferences and journals separately; Table 2 lists the distribution of invisible work for conferences and journals using the three services. Over 80% of the conference

Table 1. Mean (and standard deviation) of the invisible work rate for all areas and all services.

	WoS		WoS-P		Scopus	
Physics	0.23	(0.22)	0.15	(0.25)	0.13	(0.22)
Math	0.34	(0.22)	0.32	(0.23)	0.29	(0.23)
EE	0.71	(0.12)	0.31	(0.16)	0.25	(0.13)
CS	0.66	(0.29)	0.46	(0.20)	0.33	(0.23)

Table 2. Mean (and standard deviation) of the invisible work rate for different venues within CS.

	WoS		WoS-P		Scopus	
Conferences	0.82	(0.13)	0.47	(0.30)	0.32	(0.31)
Journals	0.32	(0.12)	0.32	(0.12)	0.16	(0.17)

publications of CS researchers were not represented in the standard WoS (with little variation among the researchers). The invisible rate drops significantly (paired t-test, p-value = 0.0003920) when using WoS-P, but the further decrease using Scopus is not significant (paired t-test p-value = 0.1512). It should be noted that the sample size is only 10, so the difference might also be significant with a larger sample size. Note, too, that almost 33% of journal publications of the 10 CS researchers was not accounted for in WoS.

Subareas of CS. We grouped the 50 CS researchers into seven subareas: artificial intelligence, architecture, interdisciplinary research, security, systems, theory, and visual computing. Interdisciplinary research involves the interface between computing and the natural sciences (biology, chemistry, and physics); security involves mainly cryptography; systems is a broad area involving distributed systems and networks; and visual computing combines computer vision and computer graphics. We classified each researcher into only a single group, but seven researchers could not be classified into any of subarea. Table 3 summarizes the number of researchers, the mean, and the standard deviation of the invisible work rate for the researchers in each subarea, for WoS, WoS-P, and Scopus, respectively.

Any analysis using subarea data can identify at most possibly interesting phenomena. The number of researchers in each group is too small for any meaningful statistical analysis and too dependent on our subjective choices of research-group associations. The data shows the relatively younger CS subarea of interdisciplinary research differs from the other CS subareas, with its researchers apparently publishing much more in journals, particularly in WoS-indexed journals, than the rest of CS.

The theory subarea, perhaps due to its proximity to math, also publishes more in journals and is particularly homogeneous regarding its publication practices. Other interesting subgroups are architecture researchers and to a lesser extent visual-computing researchers, who seem to form cohesive communities regarding publication practices. Finally, the subareas of systems and security show both high values of invisible work and high variance, indicating that on average publications of members are not well-represented in the bibliographic services and the community as a whole lacks uniform publication practices.

Researcher position. We also calculated invisible work rate based on CS researcher positions as assistant, associate, and full professor, though we had no information for three researchers. Table 4 lists mean and standard deviation of the invisible work rate for WoS, WoS-P, and Scopus, respectively, with no significant difference of invisible work rate (ANOVA p-value=0.854 for WoS, p-value=0.4492 for WoS-P, and p-value= 0.1651 for Scopus). The difference in publication practices among researchers in different stages of their careers does not seem to be significant.

Table 3. Mean (and standard deviation) of the invisible work rate for different subareas within CS.

	Number	WoS		WoS-P		Scopus	
AI	4	0.76	(0.14)	0.64	(0.21)	0.36	(0.18)
Architecture	4	0.80	(0.05)	0.45	(0.10)	0.33	(0.12)
Interdisciplinary	4	0.22	(0.24)	0.14	(0.20)	0.16	(0.21)
Security	6	0.68	(0.11)	0.47	(0.19)	0.52	(0.22)
System	14	0.72	(0.13)	0.44	(0.20)	0.40	(0.25)
Theory	7	0.54	(0.12)	0.41	(0.09)	0.13	(0.05)
Visual	4	0.69	(0.16)	0.42	(0.15)	0.31	(0.04)

Table 4. Mean (and standard deviation) of the invisible work rate for different positions within CS.

	Number	WoS		WoS-P		Scopus	
Assistant	15	0.59	(0.19)	0.42	(0.19)	0.44	(0.24)
Associate	16	0.68	(0.21)	0.52	(0.22)	0.31	(0.18)
Professor	16	0.65	(0.14)	0.42	(0.17)	0.27	(0.24)

Conclusion

When CS is classified as a science (as it was in the *U.S. News & World Report* survey), the standard bibliometric evaluations are unfair to CS as a whole. On average, 66% of the published work of a computer scientist is not accounted for in the standard WoS indexing service, a much higher rate than for scientists in math and physics. Using the new conference-proceedings service from WoS, the average invisible work rate for CS is 46%, which is higher than for the other areas of scientific research. Using Scopus, the average rate is 33%, which is higher than for both EE and physics.

CS researchers' practice of publishing in conference proceedings is an important aspect of the invisible work rate of CS. On average, 82% of conference publications are not indexed in WoS compared to 47% not indexed in WoS-P and 32% not indexed in Scopus.

The results regarding the CS subareas are tentative, pointing to results needing further investigation, including the possible homogeneity of publishing practices in the architecture, theory, and visual computing subgroups, as indicated by the low variance of the invisible work-rate distributions for these subareas. Future research should also include more researchers in each subgroup in order to identify more subtle differences in their publication practices.

Using data from Scopus, as in Table 3, we were able to verify that the subareas of visual computing and theory disagree on invisible rate, but each is consistent (low variance) within itself. Security, on the other hand, has a high variance, reflecting the tendency of the publishing practice of its researchers to vary. This heterogeneity between the subareas and within each subarea indicates it will be difficult for computer scientists to agree on a single bibliometric evaluation criterion:, any criterion is likely to generate tension between different subareas of CS and may even generate tension within a single subarea.

The similarity of the invisible rate across academic positions indicates that the research area of CS faculty members defines their publication practices more than the stage of their career.

The strength of our conclusions depends on the sample's accurate representation of the entire population of researchers in the four scientific areas and of the researchers in CS.

Our sample also involved a major source of potential bias. All researchers in the sample worked in the U.S., so some of our conclusions may not be valid elsewhere. In other countries, university and funding agency reward-and-evaluation structures may explicitly encourage researchers to concentrate on publishing their results in journals or conferences, so the invisible work rate could be different.

This research deals with the invisible rate in the two main bibliographic services from a production point-of-view; the invisible rate quantifies the number of a CS researcher's publications not accounted for in these services. However, the invisible rate in production spills over to citation count, which is the more used metric in science. CS researchers may publish papers that are well cited by others, but the papers themselves may not be included in their citation counts because they are not included in the bibliographic servers. Alternatively, papers may be included in bibliographic servers, but the number of citations to them may be severely underestimated, because many CS publications are not included in citation-count sources.

Other bibliographic services are specific to CS, including DBLP[f], the ACM Digital Library[g], and CiteSeerX[h] so cannot be used to compare CS with other scientific areas. There is also Google Scholar[i], which seems to be a general bibliographic service indexing both production and citations; we didn't include it here because it is not yet adopted (as far as we know) as the bibliographic service used for institutional metrics, probably due to the lack of transparency as to how it collects and processes bibliographic data; for a review of some of the criticism of Google Scholar as a bibliometric tool, see Bar-Ilan.[1]

Faced with multidisciplinary evaluation criteria, computer scientists should lobby for WoS-P, or better, Scopus. Understanding the limitations of the bibliometric services will help a multidisciplinary committee better evaluate CS researchers.

f http://www.informatik.uni-trier.de/~ley/db/
g http://portal.acm.org/dl.cfm
h http://citeseerx.ist.psu.edu/
i http://scholar.google.com/

On average, 66% of the published work of a computer scientist is not accounted for in the standard WoS indexing service, a much higher rate than for scientists in math and physics.

Acknowledgments

We thank Carlos J.P. Lucena and Bertrand Meyer for their support at different points in this research. We are deeply indebted to the colleague who gave us access to the WoS-C service. We also owe much to the comments, criticism, and suggestions of the two *Communications* reviewers and associate editor. The authors were supported by grants from CNPq and FAPESP.

References

1. Bar-Ilan, J. Which h-index? A comparison of WoS, Scopus, and Google Scholar. *Scientometrics 74*, 2 (Feb. 2008), 257–271.
2. Ma, R., Ni, C., and Qiu, J. Scientific research competitiveness of world universities in computer science. *Scientometrics 76*, 2 (Aug. 2008), 245–260.
3. Meho, L. and Yang, K. Impact of data sources on citation counts and rankings of LIS faculty: Web of Science vs. Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology 58*, 13 (Nov. 2007), 2105–2125.
4. Meyer, B., Choppy, C., Staunstrup, J., and van Leeuwen, J. Research evaluation for computer science. *Commun. ACM 54*, 4 (Apr. 2009), 31–34.
5. Torres-Salinas, D., Lopez-Cozar, E., and Jimenez-Contreras, E. Ranking of departments and researchers within a university using two different databases: Web of Science vs. Scopus. *Scientometrics 80*, 3 (Sept. 2009), 761–774.
6. Vardi, M.Y. Conferences vs. journals in computing research. *Commun. ACM 52*, 5 (May 2009), 5.
7. Wainer, J., Xavier, E., and Bezerra, F. Scientific production in computer science: A comparative study of Brazil and other countries. *Scientometrics 81*, 2 (Nov. 2009), 535–547.

Jacques Wainer (wainer@ic.unicamp.br) is an associate professor in the Computing Institute of the University of Campinas, Campinas, SP, Brazil.

Cleo Billa (cleo.billa@gmail.com) is an assistant professor in the Federal University of the Pampas, Alegrete, RS, Brazil.

Siome Goldenstein (siome@ic.unicamp.br) is an associate professor in the Computing Institute of the University of Campinas, Campinas, SP, Brazil.