SIDEBAR

Bibliometric Data Filters

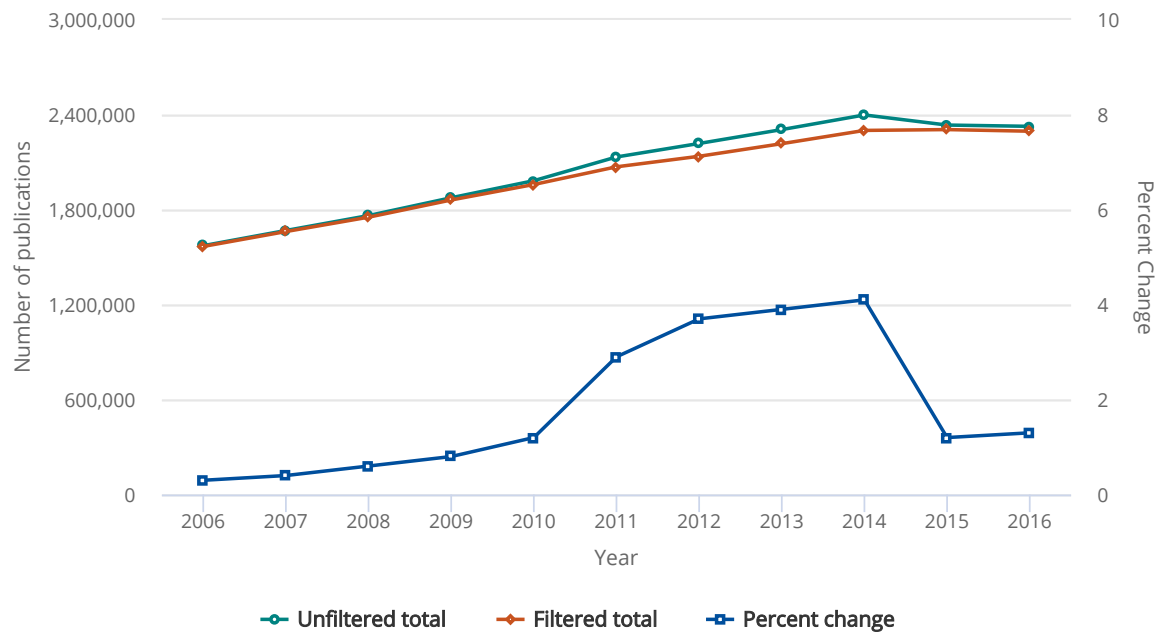
The goal of the bibliometric data analysis presented in this chapter is to measure valid peer-reviewed research output. Recently, bibliometric experts noted an increase of low-quality publications, including journals, conference proceedings, or books lacking substantive peer review.* NCSES removed two publication sets from the Scopus database to exclude low-quality publications from the bibliometric data included in this report:

- Journals and proceedings flagged by the Directory of Open Access Journals (DOAJ) for failing to adhere to its list of best practices or being suspected of editorial misconduct.†
- Elsevier's list of titles that they removed from the Scopus database from 2014 onward are removed retroactively for the *Indicators* database for all publication years.‡

The need for NCSES filtering has increased in recent years. Figure 5-D shows that the number of publications removed was 1% or less for most years, then approached 3% (more than 60,000 publications) in 2011, and grew beyond 3% (81,000–98,000 publications) each year from 2012 to 2014. The number of publications filtered for the *Indicators* database dropped back down to the 1% range in 2015–16 as Elsevier began instituting filters on the Scopus database.

FIGURE 5-D

Filtered and unfiltered publications in Scopus, by year: 2006–16



Note(s)

Percent change is computed as the difference of publications between the filtered and the unfiltered approaches divided by the number of publications in the unfiltered approach.

Source(s)

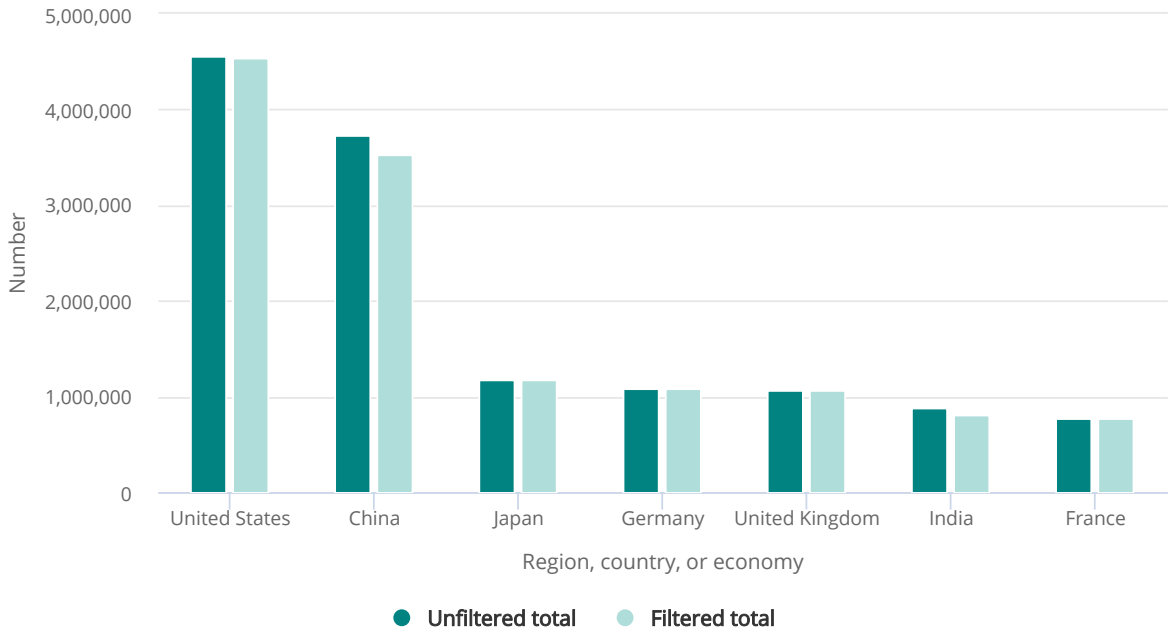
Science-Metrix; Elsevier, Scopus abstract and citation database, accessed July 2017.

Science and Engineering Indicators 2018

Figure 5-E shows the numerical impact of the filters by country or economy. During the last 11 years for which data are available, 2006–16, China saw the most removed publications (more than 215,000 publications removed; approximately 6% of its publication total; more than 50% of all removed publications), followed by India (nearly 62,000 publications removed; 7.6% of its publication total; nearly 14% of all removed publications). Other countries or economies notably affected by this filtering (but not shown in Figure 5-E) include Iran and Malaysia, each had approximately 18,000 publications removed. In the case of Malaysia, this accounted for more than 13% of its total publication output. Beyond these, only Russia and South Korea had more than 8,000 publications removed (about 2% of all publications removed each).

FIGURE 5-E

Filtered and unfiltered publications in Scopus, by region, country, or economy: 2006–16



Source(s)

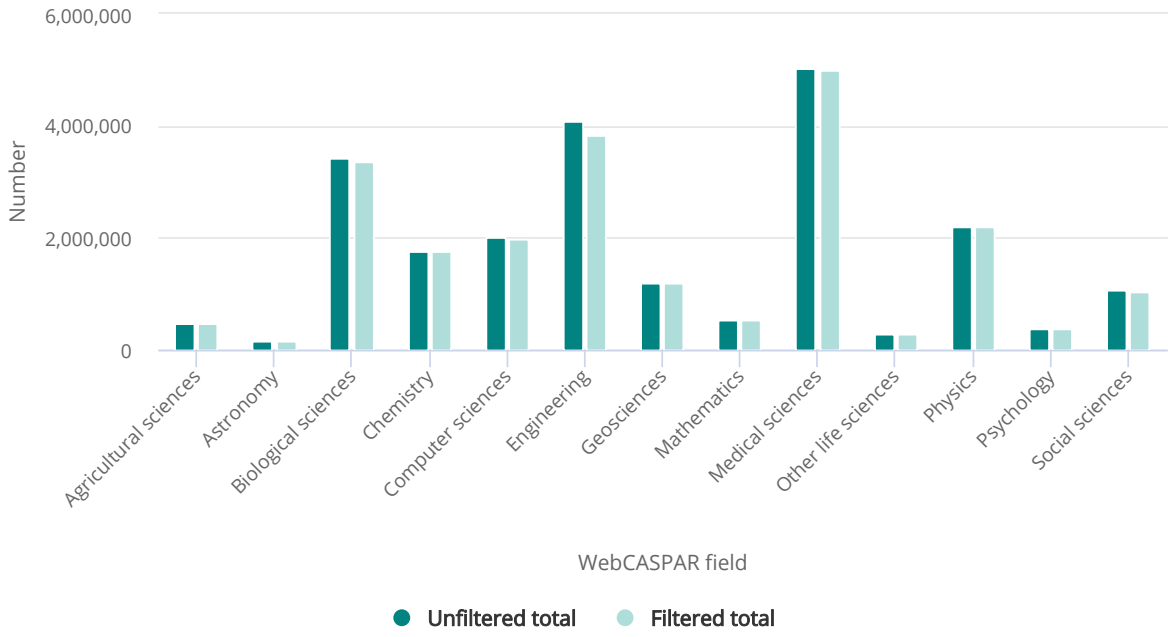
Science-Metrix; Elsevier, Scopus abstract and citation database, accessed July 2017.

Science and Engineering Indicators 2018

The majority of removed publications are conference proceedings. For example, cases where publishers post new conference proceedings every day, each containing many articles, sends a clear red flag concerning robustness, originality, and peer review (Van Noorden 2014). In addition, the biggest filter impact by field is on engineering, where more than 6% of the publications (more than 250,000) were removed in this filtering process (Figure 5-F). This is because conference proceedings comprise both a large share of the removed publications (Table 5-D) and are a large share of the engineering publications.

FIGURE 5-F

Filtered and unfiltered publications in Scopus, by WebCASPAR field: 2006–16



WebCASPAR = Integrated Science and Engineering Resources Data System.

Source(s)

National Science Foundation, National Center for Science and Engineering Statistics; Science-Metrix; Elsevier, Scopus abstract and citation database, accessed July 2017.

Science and Engineering Indicators 2018

TABLE 5-D

Number of titles and publications filtered from the Scopus database

(Number)

Filter	Journals		Conference proceedings		Total	
	Titles	Publications	Titles	Publications	Titles	Publications
Scopus	216	162,323	15	238,208	231	400,531
Directory of Open Access Journals	106	67,497	0	0	106	67,497
Total	307	211,595	15	238,208	322	449,803

Note(s)

"Titles" includes journals, books, and conference proceedings, and "Publications" includes the individual items appearing in the titles. Prepared by Science-Metrix using Scopus (Elsevier). Total does not sum to individual sources because there is some overlap between Directory of Open Access Journals and Scopus filters.

Source(s)

Science-Metrix; Elsevier, Scopus abstract and citation database, accessed July 2017.

Science and Engineering Indicators 2018

* For an example of journals requiring robust and novel submissions, see https://www.nature.com/authors/policies/peer_review.html. Articles on predatory publication are https://www.nytimes.com/2016/12/29/upshot/fake-academe-looking-much-like-the-real-thing.html?_r=0, https://www.nytimes.com/2013/04/08/health/for-scientists-an-exploding-world-of-pseudo-academia.html, http://science.sciencemag.org/content/342/6154/60.full, and https://www.nature.com/news/predatory-publishers-are-corrupting-open-access-1.11385.

† The DOAJ list of excluded journals is available at https://docs.google.com/spreadsheets/d/183mRBRqs2jOyP0qZWXN8dUd02D4vL0Mov_kgYF8HORM/edit. Note that DOAJ also flags serials that are no longer available in open access (OA); although an important and evolving phenomenon in the research landscape, OA status is not associated here with any specific demarcation of quality, whether low or high. Thus, the titles flagged by DOAJ for OA-related reasons alone are not filtered out of the database for *Science and Engineering Indicators 2018*.

‡ Elsevier's principles of quality can be found at https://www.elsevier.com/solutions/scopus/content/content-policy-and-selection and https://doaj.org/bestpractice. In 2014, during its periodic reevaluation of items flagged for follow-up, the Scopus Content Selection and Advisory Board elected to remove 42 titles as of 2014. The 42 titles are retroactively removed from the *Indicators* database to create a valid time series for bibliometric analysis, even though Elsevier does not claim that these titles were necessarily of low quality before 2014.