

Duplication issues with the new interface of *Scopus*

Bakthavachalam Elango

Como citar este artículo:

Elango, Bakthavachalam (2024). "Duplication issues with the new interface of *Scopus*". *Infonomy*, 2(1) e24015.

<https://doi.org/10.3145/infonomy.24.015>



Bakthavachalam Elango

<https://orcid.org/0000-0002-8938-0155>

<https://www.directorioexit.info/ficha6493>

Rajagiri College of Social Sciences

Department of Library and Information Science

& Centre of Excellence in Journalology

Kochi 683 104 Kerala, India

elangokb@yahoo.com

Funding: None

Conflict of interest: The author has no conflict of interest to disclose.

Use of AI text: To improve the phrasing and readability, writing tools such as *QuillBot* were used.

Abstract

This short communication discusses the duplication issue in *Scopus*'s new interface. After searching for bibliographic records, the author noticed many duplicates in the downloadable file. This issue lasted a month, with 67% of records duplicated. After alerting *Scopus*, the new version was confirmed to have a duplication issue, prompting a recommendation to revert back to the old version. The study advises caution when utilizing *Scopus* data, particularly when exporting large files. Additionally, it highlights the novel attributes of the *Scopus* interface, including the ability to download 20,000 records at a time, a notable increase from the previous limit of 2,000.

Keywords

Scopus; Bibliometrics; Bibliographic records; Duplication; Journalology; File import; File export; Downloading files.

1. Introduction

Scopus, an abstract and indexing database with full-text links, was introduced by *Elsevier* in 2004, as an alternate to the *Web of Science*. Its name, *Scopus*, was derived from the bird known as Hammerkop (*Scopus umbretta*), which is reputed to its exceptional navigational abilities (Burnham, 2006). Among the vast array of curated abstract and citation databases, *Scopus* stands as one of the largest, encompassing a comprehensive range of scientific journals, conference proceedings, and books on a global scale. Its commitment to ensuring the highest quality of indexed data by an independent *Content Selection and Advisory Board* (Baas et al., 2020). Since its inception, numerous studies undertaken to evaluate *Scopus*'s coverage (Singh et al., 2021; Pranckutė, 2021; Mongeon & Paul-Hus, 2016; Vera-Baceta et al., 2019). Most importantly, *Scopus* maintains an ongoing assessment of journals submitted by publishers for indexation, ensuring the continual verification of quality (Krauskopf, 2018).

A study conducted by Thelwall (2018) found that all *Scopus* articles with DOIs were found in *Dimensions* (97% in 2012) and claimed that the *Dimensions* proved to a viable alternative to *Scopus* and *Web of Science* for general citation analyses and citation support. In a recent assessment, Thelwall & Sud (2022) examine the changes in the extent of coverage provided by *Scopus*, a prominent citation index, over a span of 121 years starting from 1900. Since *Scopus*'s coverage is limited to a fraction of journal publications outside of Europe and North America, and as such, it cannot be classified as a truly global database (Tennant, 2020).

Despite facing some criticism, the *Scopus* database is employed not solely for bibliometric evaluations (Elango et al., 2019; Elango et al., 2023), but also for the purpose of ranking frameworks on a global scale, such as the *Times Higher Education World University Rankings*, as well as on a national level, such as the *National Institutional Ranking Framework* in India. In addition to this, some researchers have integrated the *Scopus* data into their respective research studies (Ceasar & Ignacimuthu, 2023).

2. New *Scopus* interface

Scopus has implemented a trial version of its new interface as of August 2022, which boasts a range of novel features. Notably, users are now able to download up to 20,000 records or first 20,000 records at a time, a substantial increase from the previous limit of 2,000 records in the older version.

Our college library has subscribed to the *Scopus* database. In the second week of December 2023, we undertook a retrieval of pertinent bibliographic records using the following search string:

ethnobotan* OR ethnove* OR ethnomed* OR "traditional knowledge" OR "alternative medicine" OR "Herbal medicine" OR "Folk Medicine"
with a particular focus on Indian affiliations.

Regrettably, we discovered a significant number of duplicate records within the downloaded file. To verify this occurrence, we contacted other subscribing institutions, who subsequently confirmed the presence of these duplicates.

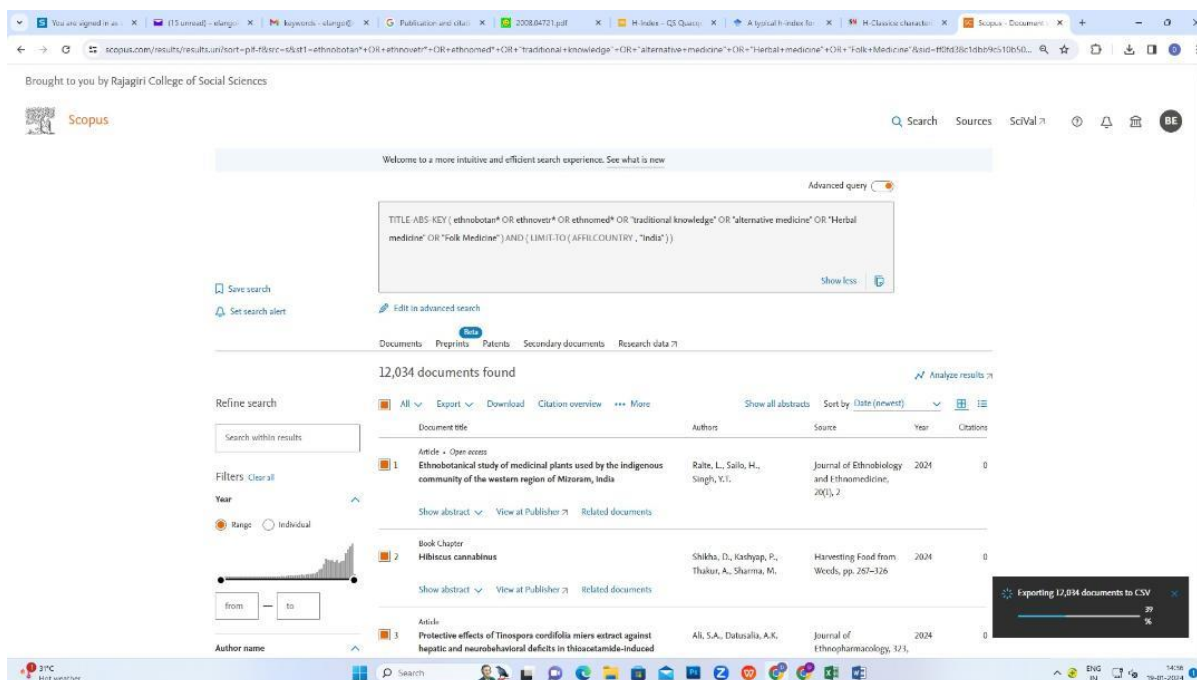


Figure 1. Search results in Scopus

Even after a month, this issue persisted until January 19, 2024. For instance, the aforementioned search parameters yielded a total of 12,034 records (see Figure 1), of which 67% were found to be duplicates (see Figure 2), indicating a two-thirds duplication rate. After informing the Scopus team of this dilemma, they acknowledged the existence a duplication problem in the new version. As a resolution, they recommended reverting back to the old version (see Figure 3). Consequently, a cautious message has been incorporated into the new interface, alerting users to the possibility of encountering missing or duplicate records when exporting large files containing more than 2,000 records (see Figure 4).

3. Remarks

In order to optimize the utilization of Scopus data, it is imperative to exercise a significant level of caution. The context under which the data is obtained bears no significance; this remains a constant reality.

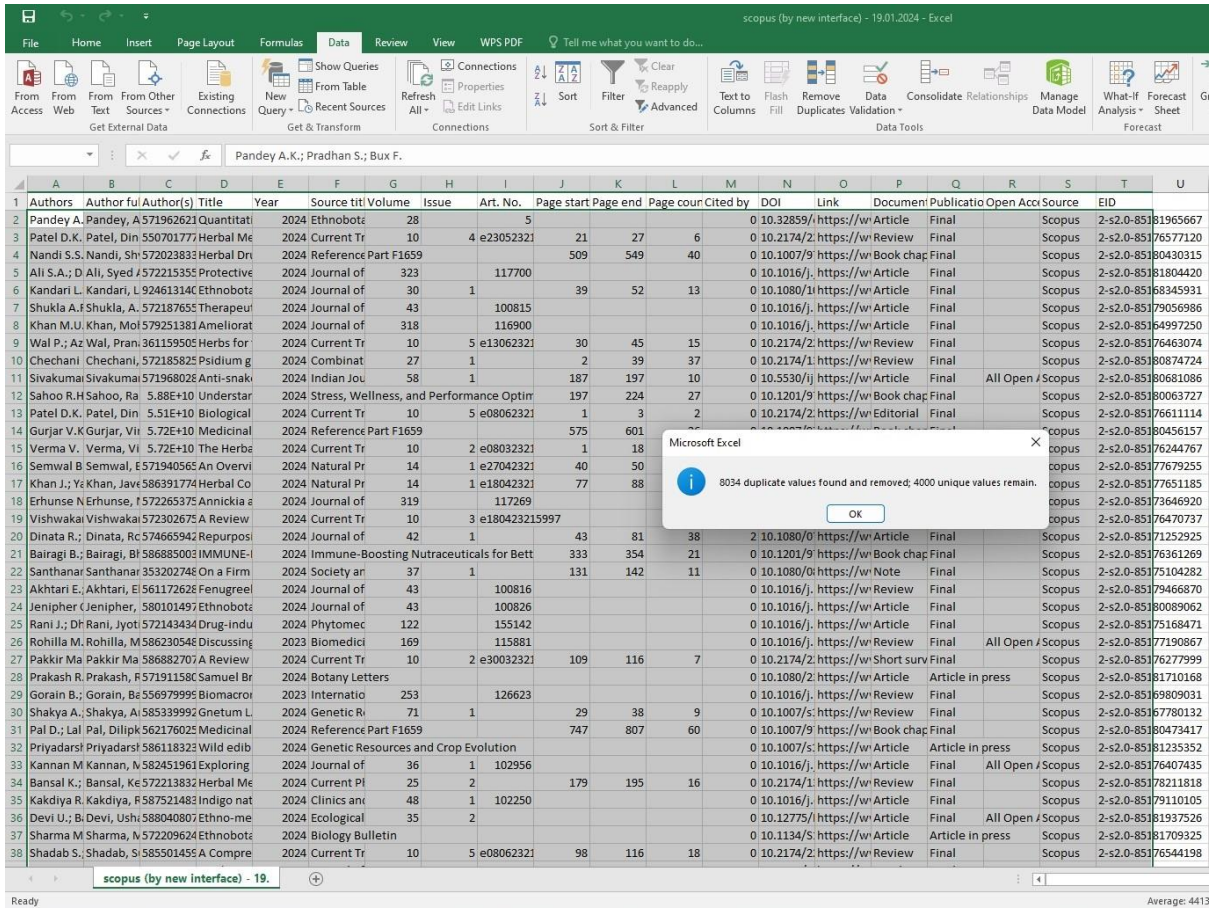


Figure 2. Showing the number of duplicate records in Excel

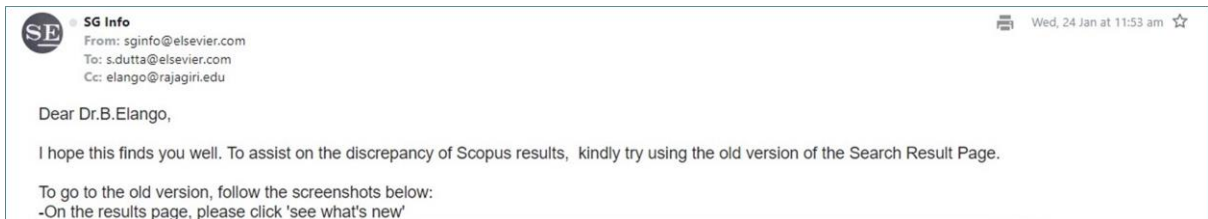


Figure 3. Communication received from Scopus team

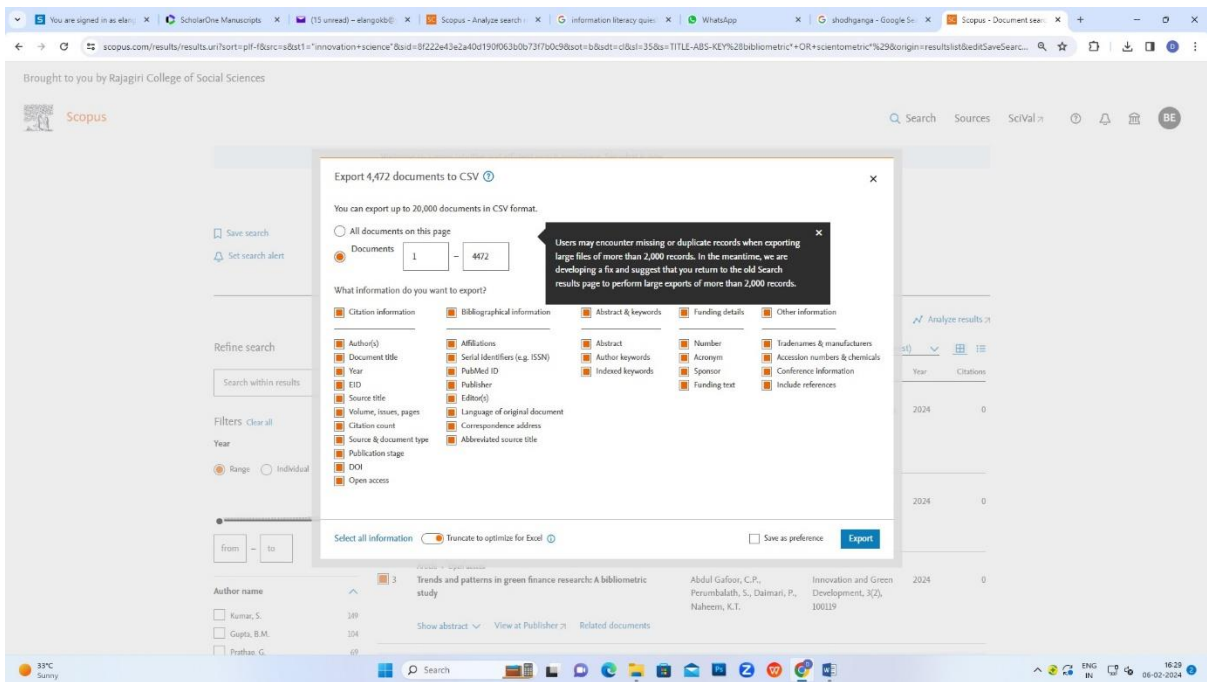


Figure 4. Caution showing the duplicate records existed

4. References

Baas, J.; Schotten, M.; Plume, A.; Côté, G.; Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative science studies*, 1(1), 377-386.

https://doi.org/10.1162/qss_a_00019

Burnham, J. F. (2006). Scopus database: a review. *Biomedical digital libraries*, 3(1), 1-8.

<https://doi.org/10.1186/1742-5581-3-1>

Cesar, S. A.; Ignacimuthu, S. (2023). CRISPR/Cas genome editing in plants: Dawn of Agrobacterium transformation for recalcitrant and transgene-free plants for future crop breeding. *Plant Physiology and Biochemistry*, 196, 724-730.

<https://doi.org/10.1016/j.plaphy.2023.02.030>

Elango, B.; Kozak, M.; Rajendran, P. (2019). Analysis of retractions in Indian science. *Scientometrics*, 119(2), 1081-1094.

<https://doi.org/10.1007/s11192-019-03079-y>

Elango, B.; Matilda, S.; Martina Jose Mary, M.; Arul Pugazhendhi, M. (2023). Mapping the cybersecurity research: A scientometric analysis of Indian publications. *Journal of Computer Information Systems*, 63(2), 293-309.

<https://doi.org/10.1080/08874417.2022.2058644>

Krauskopf, E. (2018). An analysis of discontinued journals by Scopus. *Scientometrics*, 116(3), 1805-1815.

<https://doi.org/10.1007/s11192-021-03948-5>

Mongeon, P.; Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, 106, 213-228.
<https://doi.org/10.1007/s11192-015-1765-5>

Pranckutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications*, 9(1), 12.
<https://doi.org/10.3390/publications9010012>

Singh, V. K.; Singh, P.; Karmakar, M.; Leta, J.; Mayr, P. (2021). The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. *Scientometrics*, 126, 5113-5142.
<https://doi.org/10.1007/s11192-021-03948-5>

Tennant, J. P. (2020). Web of Science and Scopus are not global databases of knowledge. *European Science Editing*, 46, e51987.
<https://doi.org/10.3897/ese.2020.e51987>

Thelwall, M. (2018). Dimensions: A competitor to Scopus and the Web of Science?. *Journal of informetrics*, 12(2), 430-435.
<https://doi.org/10.1016/j.joi.2018.03.006>

Thelwall, M.; Sud, P. (2022). Scopus 1900–2020: Growth in articles, abstracts, countries, fields, and journals. *Quantitative Science Studies*, 3(1), 37-50.
https://doi.org/10.1162/qss_a_00177

Vera-Baceta, M. A.; Thelwall, M.; Kousha, K. (2019). Web of Science and Scopus language coverage. *Scientometrics*, 121(3), 1803-1813.
<https://doi.org/10.1007/s11192-019-03264-z>