A user-friendly method to merge Scopus and Web of Science data during bibliometric analysis

Accepted to be published in ***Journal of Marketing Analytics***

Accepted version
https://doi.org/10.1057/s41270-021-00142-7

Andrea Caputo, PhD (*corresponding author*)
Department of Economics & Management, University of Trento, Trento, Italy
&
Department of Management, University of Lincoln, Lincoln, United Kingdom

Mariya Kargina
Department of Management & Law, University of Rome "Tor Vergata", Rome, Italy

Abstract

Bibliometric studies in management and related fields are growing exponentially due to the need to systematize and summarize the growing body of publications. To do so, scholars mostly retrieve publications and metadata from either Scopus or Web of Science. Only a few bibliometric studies merge the two databases to conduct a single integrated analysis, but they do not specify how they did it. Recent studies demonstrated the benefits of merging data from Scopus and Web of Science and presented methods for the merging. In this paper we build upon a recent method to simplify some of the key steps of merging datasets when using the R package Bibliometrix to perform bibliometric analyses. The result is a user friendly, accessible, three-step method that allows researchers to save time without compromising the integrity of the data, and the analysis. Our method is particularly beneficial for a wider application as it does not require coding skills, and neither proprietary nor shareware software.

Keywords: Bibliometric analysis; Scopus; Web of Science; WOS; Merged databases; Bibliometrix; Excel; method

Introduction

Accessibility to the research data and means are fundamentally important to ensure transparency, openness, and reproducibility of science (Nosek *et al.*, 2015). An open research world needs not only access to knowledge, but also to the methods and software packages to perform the research. This article intends to contribute to the development and use of bibliometric methods that are widely available and accessible to researchers around the worlds. In particular, building upon the work of Echchakoui (2020) and previous bibliometric studies (e.g., Caputo *et al.*, 2021) we present a user friendly and open method to merge bibliographic metadata from the Scopus and Web of Science (WOS) databases to perform bibliometric analysis via the free R package Bibliometrix (Aria and Cuccurullo, 2017b). The issue of merging metadata from different databases is important in bibliometric research

because the singularities of each database may impact the results of the analyses (Echchakoui, 2020).

Synthesizing and systematizing past research findings is among the most important tasks for advancing research (Zupic and Čater, 2015) and is needed in a world that publishes millions of articles each year at a growing rate – WOS indexed more than 45 million papers in the period 1973-2013 (Larivière, Haustein and Mongeon, 2015). In this context bibliometrics, i.e. "the quantitative study of physical published units or of bibliographic units or of the surrogates for either" (Broadus, 1987, p. 376), has received growing attention as a tool for science mapping. This has been adopted as a stand-alone technique to provide an overview picture of the structure and development of a field of research (Ferreira, 2018), or together as an instrumental step in systematic or narrative reviews to support a more detailed account a research area and its evolution (Dabić *et al.*, 2020). The adoption of bibliometric methods is increasingly popular as they introduce a systematic, transparent, and reproducible review process that limits the inherent subjectivity and potential biases of qualitative based reviews (Tranfield, Denyer and Smart, 2003; Zupic and Čater, 2015). Recent bibliometric analyses supported the systematization of several research areas in management such as business economics (Castillo-Vergara, Alvarez-Marin and Placencio-Hidalgo, 2018), international entrepreneurship (Baier-Fuentes *et al.*, 2019), international business (Dabić *et al.*, 2020), business models (Caputo *et al.*, 2021), artificial intelligence in marketing (Vlačić *et al.*, 2021), smart urban ecosystems (Palumbo *et al.*, 2021) and knowledge management (Fakhar Manesh *et al.*, 2021), among many others.

To perform bibliometric analyses, researchers search for publications primarily on electronic abstract and citation databases, primarily Elsevier's Scopus and Clarivate Analytics' Web of Science (WOS). Their coverage differs substantially, depending on the research area, and therefore the results of bibliometric analyses may vary depending on the database used (Mongeon and Paul-Hus, 2016; Echchakoui, 2020). However, most scholars either use only one of the two databases or when they use both they do so by analyzing results separately because of the difficulty to manually merge them due to the different tag fields structures (Echchakoui, 2020). For example, Web of Science considers citations received by a document as "Times Cited", while Scopus uses the tag "Cited by". Echchakoui (2020) reports how only a very few studies merge the search results from both databases and they do so without a clear description of the steps. To solve this issue, the author developed and presented a step-by-step method to merge databases via several software packages (e.g., EndNote, MS Word VBA, R Studio) and codes. Even if the method is clearly explained and useful it poses some issues for wide adoption, such the fact that some of the software used is not widely and freely accessible, i.e., EndNote; some coding in Microsoft Word VBA cannot be performed on all operative systems; not all researchers are confident and comfortable with coding and using a few different software to perform the merge. These issues make the method not entirely accessible, which may hinder its wider application in support to researchers when merging datasets.

To solve this issue, we have developed and tested a manual method of merging that can be performed quite easily and uses mostly either freely available software, i.e., R, R Studio, Bibliometrix, or widely available software in academia, i.e., Microsoft Excel. The benefits of the method are its accessibility, ease of use, and speed. Despite being manual and hence

subject to human error, the method is accurate and effective. Similarly, to Echchakoui (2020), our method has been tested to be used with the Bibliometrix package in R Studio.

Considerations on WOS, Scopus and merging methods

To analyze the state of research about a topic or research field, researchers usually adopt systematic literature review and bibliometric methods collecting previous publications from scholarly databases (Paul and Criado, 2020). In management research, whether the review effort is based on a list of top journals (Iacobucci *et al.*, 2019) or on the overall field of study (Paul and Criado, 2020) articles are mostly retrieved from well-established bibliographic databases such as Web of Science (WOS) and Scopus.

Being one of the oldest scholarly databases Web of Science (WOS) is a database owned by Clarivate Analytics. The Web of Science Core Collection is considered as the premier resource of the platform, and it includes more than 1.5 billion cited references, 254 disciplines, and 74.8 million scholarly data and datasets (*Web of Science Core Collection*, 2021). WOS was the main source of bibliometric data until the launch of Scopus by Elsevier in 2004, which had quickly become the main competitor to the prevailing role of WOS. Scopus is similarly a bibliographic multidisciplinary database and considered to have one of the largest coverages of academic journals focusing on the comprehensiveness of the incoming scholarly content while WOS makes a focus on the selectivity (Visser, van Eck and Waltman, 2021). Both databases are the most used data source platforms of academic and scientific research (Zhu and Liu, 2020), especially regarding bibliometric studies (Ellegaard, 2018; Visser, van Eck and Waltman, 2021). A number of studies evaluated WOS and Scopus databases and made comparisons from different angles (Mongeon and Paul-Hus, 2016; Abdulhayoglu and Thijs, 2018; Martín-Martín *et al.*, 2018; Zhu and Liu, 2020). For instance, the correlation rate between WOS and Scopus is perceived to be considerable on several dimensions (Archambault *et al.*, 2009): coverage overlap showed that 54% of the titles in Scopus were indexed in WOS, and 84% of the titles in WOS were found in the Scopus database (Gavel and Iselid, 2008). Thus, it is essential for conducting a valid and comprehensive bibliometric research to evaluate and select the database to use. In order to identify which database is more suitable for the particular area, years, discipline, and topic under investigation, Neuhaus and Daniel (2008) suggest to run the same search in both databases separately, and make a final decision by screening the results obtained. However, some researchers are in favor of using the merged datasets from both databases in order to accomplish a comprehensive and valid result (Echchakoui, 2020). Sánchez et al. (2017) proposed to use both databases for more complete research due to the complementary characteristics of them as WOS provides a greater coverages in years whilst Scopus includes a wider base of academic journals. The substantial difference in coverage is also emphasized by Mongeon & Paul-Hus (2016) who warned about obtaining significantly different results in the bibliometric analyses if choosing one database over another. The majority of scholars however tend to use only one database after a careful cross-validation process that entails the comparison of results (e.g., Caputo *et al.*, 2018, 2021). While the use of a single database has its merits and may be the best choice depending on the topic to be investigated and may be covered better from one or the other database, there is the need to establish easier methods to merge databases that will allow researchers to make a choice that is not biased by the complexity of the methods, but purely by the research aim.

To address this issue Echchakoui (2020) proposed a method consisting of four steps to merge Scopus and WOS datasets and remove duplicates. The method requires the combination of several software, Endnote, R, Word and Excel VBA macros. The first step lies in converting WOS and Scopus datasets to ".bib" files from EndNote. The second step is converting the ".bib" files to "bibtex" files using R (or R Studio) software. The next step is to format the "bibtex" files to have the same tag fields in both files in Word VBA. The fourth step consists of merging the datasets and converting the files to the .xlsx formats in R (or R Studio) software. Subsequently, it is necessary to manually merge datasets and to remove the duplicates using VBA code in Excel. The implementation of the method described above could potentially include a need to possess coding knowledge and skills that are not readily available within the skill sets of many researchers in management. Despite the templates provided by the author it requires the user to have the basic coding techniques, which reduces the potential circle of users. Secondly the EndNote package does not have permanent full free access and that complicates the implementation of the Step 1 for those who do not have a subscription. These characteristics may make the use and accessibility of the method limited.

To contribute to the development of open and accessible methods to merge datasets from Scopus and WOS we build on the Echchakoui (2020) method and provide a simplified protocol that avoids the reliance on Endnote and on the coding procedures. By doing this we allow a wider audience of researchers to consider adopting merged datasets when doing bibliometric analysis.

Recommendations before merging datasets

Before proceeding to the merge, the datasets from the search results in Scopus and WOS must be created and exported. Our recommendation is to use the advanced search feature available in both databases. On both website it is possible to access the knowledge base necessary to use the Boolean operators for searching and the tag fields adopted by each database to perform the desired search. In this article we do not cover a specific search but invite consulting articles such as Echchakoui (2020) or Caputo et al. (2021) for further methodological advices about systematic search via Scopus and/or WOS for bibliometric analysis.

Once the search via carefully selected and validated keywords is performed (Tranfield, Denyer and Smart, 2003) we recommend the use of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) by Moher et al. (2009) to identify and perform the so-called "cleaning" steps to ensure the consistency and accuracy of the data being exported. PRISMA includes four steps: identification, screening, eligibility and inclusion. The Identification step concerns the database searching (in our case both Scopus and WOS) and the identification of a first dataset usually resulting from the mere application of a search string. The Screening step concerns the removal of duplicates and a first screening to exclude documents that did not meet the inclusion/exclusion criteria set up by the researcher such as language, peer review, research area, etc. The Eligibility step involves reading of the abstracts and full articles to further assess whether a record should be included or excluded from the analysis. The Inclusion step involves the export and analysis of the final datasets of publications.

It is advisable that a panel of experts or a group of researchers perform these steps independently to ensure the internal validity of the search (Caputo *et al.*, 2016). In the case of a single researcher, the panel of experts may be formed externally to the project (Caputo, 2013). Moreover, technological updates in both Scopus and WOS allow now for an easier use of the features of saved searches and lists. The researcher, instead of exporting data outside of Scopus or WOS in the Identification or Screening steps, can also perform the Eligibility step in the databases via the use of the saved lists.

Having reached the Eligibility step the researcher will export from Scopus and WOS the results in ".bib" (Step 1 from Table 1). The following section provides the details of the method for merging the datasets from Scopus and WOS.

Merging Scopus and WOS datasets with Bibliometrix and Excel for bibliometric analysis

To merge the two datasets retrieved from Scopus and WOS we propose the following three-step procedures that in our opinion simplify it without compromising the outcome of the method outlined by Echchakoui (2020). The method we developed for merging Scopus and WOS datasets has been tested and works for bibliometric analysis run with Bibliometrix. However, variations may be made to the method to be used with other bibliometric software such as VOSViewer (Van Eck and Waltman, 2010). The method also uses Microsoft Excel, but we expect it to perform the same steps in alternative software such as Apple Numbers or OpenOffice Calc.

Bibliometrix (Aria and Cuccurullo, 2017b) is an open source software whose sources are published on the official "*Comprehensive R Archive Network*" (CRAN) and on GitHub. The website www.bibliometrix.org provides information about the features, documents, download, FAQ, and publications about the software. The Bibliometrix software has the biblioshiny app, which provides a web-interface for Bibliometrix supporting scholars in an easy use of the main features of Bibliometrix without coding.

Microsoft Excel is a spreadsheet software for data analysis and documentation that contains a number of columns and rows where each intersection of a column and a row is a "cell." Each cell contains one point of data or one piece of information. The program is particularly useful in bibliometric analysis due to its versatility in handling files exported from Scopus and Web of Sciences such as ".csv".

--- please insert table 1 about here ---

Step 1: Convert Scopus and WOS datasets to bibliography files.
To convert Scopus and WOS datasets to bibliography files is sufficient to export each dataset in the ".bib" file format. In WOS this is possible either from the search results or marked list page by selecting "Export" and then "Other file formats". At this point the researcher will select "Full Record and Cited References" from the Record Content drop down menu and "BibTex" from the File Format drop down menu. Please note that WOS allows export of up to 500 records at the time. If the search is larger the researcher can proceed with multiple

exports and then copy paste them into a single ".bib" file by using a simple notepad or text editor program.

In Scopus either from the search result or the saved list page, the researcher can proceed by first ticking all records, the "Select All" function and then select the "BibTex" export making sure to select the necessary information to export such as references, keywords, and citation information, depending on the type of analyses to perform. We suggest saving each ".bib" as either Scopus.bib or WOS.bib for ease of following the next steps.

Step 2: Convert both Scopus and WOS .bib files to Bibliometrix format.
Having retrieved the two ".bib" files for each database search the researcher can proceed to converting them into bilbliometrix format. To do so we used the Bibliometrix package in R Studio through the "biblioshiny()" command that allows the use of a user friendly web interface (Aria and Cuccurullo, 2017a). Using biblioshiny allows the step to be also performed by a researcher not familiar with using R Studio.

The procedure to convert the files is the same and needs just to be repeated for each of them. The researcher will load the ".bib" file in Bibliometrix via the "Import raw files" option in the "Import or Load files" menu under "Data". Then, the researcher will save as "Excel" from the "Export a bibliometrix file" menu in the same page. We suggest saving each file as Scopus.xlsx and WOS.xlsx for ease of process.

This step has now created two Excel files that have the same tag fields in the first row, but information will be arranged in a different order of the column. We have also noticed that some information that is unnecessary for bibliometric analysis such as the URL link to the paper may be missing in the WOS dataset.

Step 3: Merge the two datasets manually from Excel.
The third and final step concerns the actual merging of the dataset that the researcher now has in the same format. To do this, open both files in Excel and arrange the windows to horizontal view and this will allow the researcher to have the two excel files one above the other on the screen. Then, choose one of the two as the master file to merge into from the other, for example Scopus.xlsx. In the chosen master file remove all columns that pertain to information that is not necessary for the analysis (see the Bibliometrix manual and support files for the description of the tag fields). In the second file manually remove the column pertaining to unnecessary information and rearrange the order of the remaining column e.g., WOS.xlsx to be identical to the order of the column in the master file. Having obtained two excel files with identical structure you can proceed to copy and paste the content of one into the other to obtain a merged dataset. Finally, duplicates can be removed by using the "Remove Duplicates" function under the Data menu in Excel.

Having completed these three simple steps the researcher will have obtained one Excel file that merged the results from both Scopus and WOS searches that can be loaded into Bibliometrix, this time as "Load bibliometrix file(s)" to perform the bibliometric analyses.

Concluding considerations

This paper presented a user friendly and accessible method to merge dataset retrieved from Scopus and WOS search results to perform bibliometric analyses. It does so by building upon the method developed by Echchakoui (2020).

Although the paper has several methodological and practical contributions for researchers who adopt bibliometric analyses on several fields our considerations mostly come from the application of bibliometric research on areas related to management and business at large.

The proposed three-steps method is user friendly to use and easily accessible both in terms of access to software and adoption of them as it does not require coding skills nor advanced use of R Studio beyond the Bibliometrix package necessary for bibliometric analyses. The method also does not require the use of shareware software making it accessible to potentially all researchers, especially those from less funded institutions. In addition, the steps that we adopted with Microsoft Excel could be performed with alternatives such as OpenOffice Calc.

It is worth noting that the method may be subject to human error as every manual step in bibliometric analysis tends to do, although we believe the ease of use and the benefits are superior to the risk of committing a manual error. In particular, the method allows the researcher to view and familiarize themselves with the datasets, a feature that can help in spotting exporting issues such as truncated rows or missing information that may otherwise remain hidden when adopting other more automatic methods of merging.

A point to be raised that could be the subject of further debate and methodological innovation is about the comparability of citation metadata when merging datasets from Scopus and WOS. In fact, both Scopus and WOS are closed systems that count citations obtained by the documents in the database from citing documents present in the same database. This means that the same record may have different citation counts depending on where the data is retrieved from. As both Scopus and WOS have their singularities, advantages, and disadvantages that we did not discuss in detail in this article, we note that it is the researcher's role to make a careful and solid choice to decide which of the two databases is primary even when merging. Similarly to the studies that justify the adoption of only one database, it is also recommended that those who merge databases follow a cross-validation procedure to evaluate which database to consider primarily (e.g., Caputo *et al.*, 2021). This cross-validation should not only be driven by the data of the specific search but also from existing research on the strength and weaknesses of Scopus and WOS, depending on the scope and focus of the subject under investigation (e.g., Visser, van Eck and Waltman, 2021).

Authors biographies:

Andrea Caputo is Associate Professor at the University of Trento (Italy) and at the University of Lincoln (UK). He received his PhD in Management from the University of Rome Tor Vergata, Italy. His main research interests are in entrepreneurship, negotiation, decision-making, internationalization, and strategic management. He published in several international journals, including Human Resource Management J, J of Business Research, Small Business Economics, and J of Cleaner Production among the others. Andrea is Associate Editor of the

Journal of Management & Organization and Editor of the Emerald Book Series "Entrepreneurial Behavior".

Mariya Kargina is a PhD student in Organizational Behavior at the Department of Management and Law of the University of Rome "Tor Vergata", Italy. She received her Master of Science from the University of Lincoln (UK). Her research interests are focused on international management, cross-cultural management, virtuality in team management and cultural intelligence.

Conflict of interest: The authors declare that they have no conflict of interest.

References

Abdulhayoglu, M. A. and Thijs, B. (2018) 'Use of locality sensitive hashing (LSH) algorithm to match Web of Science and Scopus', *Scientometrics*. Springer, 116(2), pp. 1229–1245.

Archambault, É. *et al.* (2009) 'Comparing bibliometric statistics obtained from the Web of Science and Scopus', *Journal of the American society for information science and technology*. Wiley Online Library, 60(7), pp. 1320–1326.

Aria, M. and Cuccurullo, C. (2017a) 'A brief introduction to bibliometrix', *Journal of Informetrics*, 11(4), pp. 959–975.

Aria, M. and Cuccurullo, C. (2017b) 'Bibliometrix: An R-tool for comprehensive science mapping analysis', *Journal of Informetrics*. Elsevier Ltd, 11(4), pp. 959–975. doi: 10.1016/j.joi.2017.08.007.

Baier-Fuentes, H. *et al.* (2019) 'International entrepreneurship: a bibliometric overview', *International Entrepreneurship and Management Journal*. Springer, 15(2), pp. 385–429. doi: 10.1007/s11365-017-0487-y.

Broadus, R. N. (1987) 'Toward a definition of "bibliometrics"', *Scientometrics*, 12(5–6), pp. 373–379. doi: 10.1007/BF02016680.

Caputo, A. (2013) 'A Literature Review of Cognitive Biases in Negotiation Processes', *International Journal of Conflict Management*, 24(4), pp. 274–398. doi: 10.1108/IJCMA-08-2012-0064.

Caputo, A. *et al.* (2016) 'Internationalisation of firms from Central and Eastern Europe', *European Business Review*. Emerald Group Publishing Limited, 28(6), pp. 630–651. doi: 10.1108/EBR-01-2016-0004.

Caputo, A. *et al.* (2018) 'Conflict management in family businesses: A bibliometric analysis and systematic literature review', *International Journal of Conflict Management*, 29(4), pp. 519–542. doi: 10.1108/IJCMA-02-2018-0027.

Caputo, A. *et al.* (2021) 'Digitalization and business models: Where are we going? A science map of the field', *Journal of Business Research*, 123, pp. 489–501. doi: 10.1016/j.jbusres.2020.09.053.

Castillo-Vergara, M., Alvarez-Marin, A. and Placencio-Hidalgo, D. (2018) 'A bibliometric analysis of creativity in the field of business economics', *Journal of Business Research*. Elsevier, 85, pp. 1–9.

Dabić, M. *et al.* (2020) 'Pathways of SME internationalization: a bibliometric and systematic review', *Small Business Economics*, 55(3). doi: 10.1007/s11187-019-00181-6.

Echchakoui, S. (2020) 'Why and how to merge Scopus and Web of Science during bibliometric analysis: the case of sales force literature from 1912 to 2019', *Journal of Marketing Analytics*. Springer, 8(3), pp. 165–184.

Van Eck, N. J. and Waltman, L. (2010) 'Software survey: VOSviewer, a computer program for bibliometric mapping', *Scientometrics*. Springer, 84(2), pp. 523–538. doi: 10.1007/s11192-009-0146-3.

Ellegaard, O. (2018) 'The application of bibliometric analysis: disciplinary and user aspects', *Scientometrics*. Springer Netherlands, 116(1), pp. 181–202. doi: 10.1007/s11192-018-2765-z.

Fakhar Manesh, M. *et al.* (2021) 'Knowledge management in the fourth industrial revolution: Mapping the literature and scoping future avenues', *IEEE Transactions on Engineering Management*. IEEE, 68(1), pp. 289–300.

Ferreira, F. A. F. (2018) 'Mapping the field of arts-based management: Bibliographic coupling and co-citation analyses', *Journal of Business Research*. Elsevier, 85(March 2017), pp.

348–357. doi: 10.1016/j.jbusres.2017.03.026.

Gavel, Y. and Iselid, L. (2008) 'Web of Science and Scopus: A journal title overlap study', *Online Information Review*. Emerald Group Publishing Limited, 32(1), pp. 8–21. doi: 10.1108/14684520810865958.

Iacobucci, D. *et al.* (2019) 'The state of marketing analytics in research and practice', *Journal of Marketing Analytics*. Springer, 7(3), pp. 152–181.

Larivière, V., Haustein, S. and Mongeon, P. (2015) 'The oligopoly of academic publishers in the digital era', *PloS one*. Public Library of Science, 10(6), p. e0127502.

Martín-Martín, A. *et al.* (2018) 'Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories', *Journal of informetrics*. Elsevier, 12(4), pp. 1160–1177.

Moher, D. *et al.* (2009) 'Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement', *PLoS medicine*. Public Library of Science, 6(7), p. e1000097.

Mongeon, P. and Paul-Hus, A. (2016) 'The journal coverage of Web of Science and Scopus: a comparative analysis', *Scientometrics*, 106(1), pp. 213–228. doi: 10.1007/s11192-015-1765-5.

Neuhaus, C. and Daniel, H. (2008) 'Data sources for performing citation analysis: an overview', *Journal of documentation*. Emerald Group Publishing Limited.

Nosek, B. A. *et al.* (2015) 'Promoting an open research culture', *Science*. American Association for the Advancement of Science, 348(6242), pp. 1422–1425.

Palumbo, R. *et al.* (2021) 'Organizing a sustainable smart urban ecosystem: Perspectives and insights from a bibliometric analysis and literature review', *Journal of Cleaner Production*. Elsevier, p. 126622.

Paul, J. and Criado, A. R. (2020) 'The art of writing literature review: What do we know and what do we need to know?', *International Business Review*. Elsevier, 29(4), p. 101717.

Sánchez, A. D., Del Río, M. de la C. and García, J. Á. (2017) 'Bibliometric analysis of publications on wine tourism in the databases Scopus and WoS', *European Research on Management and Business Economics*. Elsevier, 23(1), pp. 8–15.

Tranfield, D., Denyer, D. and Smart, P. (2003) 'Towards a methodology for developing evidence-informed management knowledge by means of systematic review', *British Journal of Management*, 14(3), pp. 207–222.

Visser, M., van Eck, N. J. and Waltman, L. (2021) 'Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic', *Quantitative Science Studies*. MIT Press - Journals, 2(1), pp. 20–41. doi: 10.1162/qss_a_00112.

Vlačić, B. *et al.* (2021) 'The evolving role of artificial intelligence in marketing: A review and research agenda', *Journal of Business Research*. Elsevier, 128, pp. 187–203.

Web of Science Core Collection (2021).

Zhu, J. and Liu, W. (2020) 'A tale of two databases: the use of Web of Science and Scopus in academic papers', *Scientometrics*, 123, pp. 321–335. doi: 10.1007/s11192-020-03387-8.

Zupic, I. and Čater, T. (2015) 'Bibliometric methods in management and organization', *Organizational Research Methods*. SAGE Publications Sage CA: Los Angeles, CA, 18(3), pp. 429–472.

Table 1 - Method to merge datasets from Scopus and WOS for bibliometric analysis with Bibliometrix

Step	Objective	Procedures	Actions
Step 1	Convert Scopus and WOS datasets to bibliography files.	Export each dataset separately directly from each database search results, in ". bib" files.	WOS: Select Export, Other File Formats, Record Content: Full Record and Cited References, File Format: BibTex Scopus: having selected All Results, select BibTex Export
Step 2	Convert both Scopus and WOS .bib files to Bibliometrix format	Use R Studio (or R) and load the Bibliometrix package, command biblioshiny() to use the web interface.	From biblioshiny: Select Import raw file(s) Import the .bib file Select Save as Excel Repeat for both Scopus and WOS .bib files
		Rename Bibliometrix Export files.	From your file manager or in Excel, rename each exported Excel file as WOS.xlsx or Scopus.xlsx
Step 3	Merge the two datasets manually from Excel	Use Excel to open both Scopus.xlsx and WOS.xldx files.	Use arrange window to horizontal for ease of use. Choose one of the two files as master file, for example Scopus.xlsx Both files will have the same tag fields in row 1, but the information will be arranged on different columns. Some less relevant information may not be present in one or the other, for example URL is present in Scopus and not in WOS.
		Simplify one of the two files, to be used as master.	Choose either Scopus.xlsx or WOS.xlsx and in Excel remove all columns that pertain to information not needed for the analysis, for example Scopus.xlsx Rearrange the column of the secondary file (for example, WOS.xlsx) in the same order of the simplified master file and remove unnecessary columns
		Copy one file to the other manually.	In Excel, paste the WOS.xlsx data to Scopus.xlsx data.
		Remove duplications.	In Excel, Use the "Remove Duplicates" function under Data menu to remove duplicate records from the merged dataset.