https://doi.org/10.51408/issi2025_076

# Document Coverage and Citation-Based Indicators: A Case Study on The Scientific Production of The Federal University of Rio De Janeiro Recovered by Web of Science, Scopus, Dimensions and Lens

Gabriel Alves Vieira[1], Jacqueline Leta[2]

[1]*gabriel.vieira@bioqmed.ufrj.br*, [2]*jleta@bioqmed.ufrj.br*
Federal University of Rio de Janeiro, Institute of Medical Biochemistry Leopoldo de Meis, Zip Code 21.941-590, Rio de Janeiro (Brazil)

## Abstract

With the growth of scientific production, quantitative indicators - such as the number of articles published in specialized journals - have assumed an increasingly central role in the evaluation of research institutions, directly influencing the allocation of resources for projects and scholarships. These indicators are directly influenced by the characteristics of the information sources used for their calculation. This study aims to investigate the impact of academic database selection on the calculation of a range of scientific output measures for a single institution: the Federal University of Rio de Janeiro (UFRJ). Four multidisciplinary bibliographic databases were selected for the retrieval of their entire set of UFRJ-related documents: Scopus, Web of Science, Dimensions and Lens. In total, 376,281 documents were retrieved and analyzed using R software. The comparative analyses performed on this corpus include an assessment of UFRJ's scientific production coverage and calculation of citation-based indicators (h, e, g, $h^c$ and i10 indices). The coverage analysis indicates a remarkably high overlap in the corpus retrieved by each source: 28% of the total documents analyzed are covered by all four sources, a percentage that increases to 36% for articles and to 49% for highly cited articles. This suggests that database size is not necessarily a critical factor in selecting an information source for scientific output analysis, especially in contexts where the focus is primarily on journal articles. Furthermore, citation-based indicators exhibited substantial variation both across databases and among the indicators themselves. Notably, a larger number of indexed documents did not necessarily correspond to higher indicator values. These findings indicate that both database choice and citation metrics selection can significantly influence the outcomes of institutional evaluation. It is therefore crucial that managers and professionals engaged in such assessments possess a thorough understanding of the characteristics and limitations of the diverse range of academic databases currently available. This knowledge is essential for selecting appropriate sources and indicators for each situation.

## Introduction

In the 1960s, with the growth of global scientific production, objective initiatives towards science evaluation became relevant to Science and Technology (S&T) managers. During this period, the OECD's Frascati manual and other instruments were developed by international bodies to promote the standardization of input and output indicators. In particular, output indicators - which measure the production of S&T documents - have increasingly assumed a central role in defining government policies (Velho, 2001), especially after the emergence of bibliographic databases focused on academic output.

The creation of academic databases - a type of secondary source that indexes metadata from ?? scientific literature (Grogan, 1970) - has significantly boosted scientometric research, which, among other objectives, aims to investigate and quantify the performance and impact of academic research (Mingers & Leydesdorff, 2015; Aria & Cuccurullo, 2017). Scientometric indicators have also been widely adopted in scientific output evaluation processes, as well as in decision-making and policy formulation by S&T managers (Mingers & Leydesdorff, 2015).

Among scientometric indicators, citation-based indicators play a particular role in the assessment of scientific production. These metrics influence not only the ranking of academic journals - often evaluated using citation-based measures (Guerrero-Bote & Moya-Anegón, 2012) - but also the advancement (or ascension) in scientific careers, as funding decisions for research projects and scholarships in many countries frequently involve evaluation processes that incorporate citation-focused indicators (Carlsson, 2009; Schneider, 2009; De Oliveira & Amaral, 2017).

A popular citation-based indicator is the h-index, which is defined as the number h of publications that have each received at least h citations (Hirsch, 2005). For example, an h-index of 25 indicates that the corpus contains 25 publications with at least 25 citations each. The set of documents that contribute to the h-index is named as the h-core, comprising the most highly cited publications within the analyzed corpus. Although originally developed to assess individual researchers, the h-index can be calculated for any collection of documents (Jones et al., 2011), making it a versatile metric for evaluating scientific output at various levels.

Over time, the h-index has inspired the development of several related indicators tailored to specific analytical needs. The *g-index* (Egghe, 2006), for example, is more sensitive to highly cited publications, while the $h^c$*-index* (Sidiropoulos et al., 2007) gives greater weight to citations received by recently published documents and the *e-index* (Zhang, 2009) differentiates h-cores based on their total citation counts.

Various indicators have also been employed in the construction of university rankings, which are typically elaborated by commercial publishers and publicized as tables that rank higher education institutions based on their performance - an assessment largely driven by quantitative data (Usher & Savino, 2009). Although these rankings are primarily targeted at the general public (such as prospective students seeking a university to attend), they also attract considerable interest within universities themselves, where they may be utilized for auditing, benchmarking, and management purposes (Johnes, 2018).

However, the use of quantitative indicators to evaluate institutional output is far from straightforward, as the choice of metrics and the weight assigned to each can significantly influence ranking outcomes, as noted by Vanz et al. (2018). Moreover, there is evidence that relying on a single database to construct these rankings can introduce bias, owing to variations in coverage across different information sources (Huang et al., 2020). Consequently, bibliographic database selection represents a critical step in the elaboration of academic rankings.

Metrics used in academic evaluation are also directly influenced by the choice of data sources (Gingras, 2016), as databases vary widely in their characteristics, structure, and coverage. For example, different databases employ distinct approaches

to document retrieval and indexing. Bibliographic databases, such as Web of Science (WoS), tend to apply strict selection criteria for the inclusion of new journals into their collections, whereas search engines, like Google Scholar, rely on web crawlers to index vast amounts of academic content available online, aiming for maximum coverage. Additionally, academic databases differ in their thematic scope: while some are multidisciplinary (e.g., Dimensions, Scopus), others specialize in specific fields, such as PubMed for the biomedical sciences or ERIC for education.

Metadata also varies across data sources. This is particularly evident in how academic disciplines are attributed to documents: databases often adopt distinct strategies for this classification, which can generally be divided into two approaches—those that assign disciplines based on the thematic scope of the publication venue, and those that classify documents directly through content analysis (Bornmann, 2018). Another field that frequently differs between sources is document type as each database typically employs its own classification scheme for categorizing the nature of the documents it indexes.

The differences among databases make their selection one of the most crucial steps in the design of any scientometric study aimed at analyzing scientific output. The growing diversity of academic databases, coupled with the need to identify the most appropriate informational source for a given purpose, has given rise to an impressive body of comparative studies examining various secondary sources. A sizable portion of the literature on bibliographic data sources focuses on comparisons between the long-established Web of Science (WoS) and Scopus databases (Archambault et al., 2009; Vieira, Gomes, 2009; Chadegani et al., 2013; Zhu, Liu, 2020). However, the introduction of new academic data platforms, including Dimensions, OpenAlex, and The Lens Scholarly Search, has prompted more recent studies to incorporate these emerging secondary sources into their comparative evaluations (Bornmann et al., 2021; Liang et al., 2021; Delgado-Quiros et al., 2023). Among the topics covered by such studies, the issue of coverage stands out as one of the most analyzed, whether at journal-level (Grindlay et al., 2012; Mongeon, Paul-Hus, 2016; Singh et al., 2021) or, more frequently, document-level (Gusenbauer, 2019; Huang et al., 2020; Martín-Martín et al., 2021; Visser et al., 2021; Gusenbauer, 2022).

The comparative analysis of citation-based indicators across various bibliographic databases serves as a valuable framework for evaluating the relationship between information sources and metrics. This approach facilitates the identification of discrepancies inherent in both the databases and the indicators themselves. Nevertheless, the existing literature on this subject is limited and outdated, frequently focused on specific disciplines and comparing a small number of databases (Franceschet, 2009). Furthermore, to our knowledge, no studies have yet explored these indicators alongside characteristics such as database coverage. Thus, this paper aims to investigate the variations in output retrieved from several databases and the impact of secondary source selection on citation-based indicators. We have opted to conduct a case study that focuses on the scientific output of a single university over its entire publication history which allows us to elucidate the effects of database selection on the assessment of institutional performance across an extended timeframe.

This work analyzes publications from the Federal University of Rio de Janeiro (UFRJ). Founded in 1920, UFRJ stands as Brazil's largest and oldest public university (Oliveira, 2019). The institution offers 176 undergraduate courses and 114 postgraduate programs (PPGs), thereby contributing to the training of professionals and the advancement of research across multiple scientific disciplines. Moreover, it ranks among the top academic institutions in Latin America, being at the 6th position in the 2024 Quacquarelli Symonds university ranking(QS, 2024) and 11th in the 2023 Times Higher Education ranking (THE, 2023). In light with its stature, UFRJ is increasingly focused on enhancing its visibility through strategic investments, including the creation of the Performance Indicator Management Office (GID - https://pr2.ufrj.br/gid), which aims to collect data for university rankings and formulate recommendations for improving the institution's classifications.

Therefore, UFRJ's relevance for Brazilian higher education and scientific development, alongside its extended publication period and increasing focus on factors influencing its standing in academic rankings, justifies its selection as our case of study. Here, we examine variations in UFRJ's scientific output across databases through two main approaches: (i) a comparative analysis of the production retrieved in multiple databases and their coverage; and (ii) an assessment of citation-based indicators calculated for each database.

## Methodology

This study was conducted in four main stages: (a) database selection; (b) data collection; (c) data processing; and (d) data analysis. These are presented in the sections below.

### *Definition of databases*

Since the reliability of our results is closely related to document retrieval accuracy, we opted against using academic search engines (e.g., Google Scholar), which tend to exhibit inconsistencies in the results yielded by the same research strategy (Gusenbauer, 2019). Considering the varied scientific output from UFRJ, it seems reasonable to assume that multidisciplinary databases are the most suitable for obtaining a representative sample of the research related to the institution.

We selected four databases: Scopus, Web of Science, Dimensions and Lens. The compatibility of the selected sources with the R bibliometrix package (Aria, Cuccurullo, 2017) was a critical factor in the selection process, as it offers the benefit of automating certain time-consuming steps of the data analysis process. Unfortunately, only these four multidisciplinary databases were supported by the package at that time.

Scopus and Web of Science are the oldest and, selective databases widely utilized in scientometric research (Baas et al., 2020; Birkle et al., 2020), whereas Lens and Dimensions are more recent databases that incorporate third-party sources (Delgado-Quirós, Ortega, 2024) and are less stringent in their indexing criteria. Thus, the selected databases also provide insights into the differences between the two distinct database models.

*Data collection*

For this stage, all documents indexed by the four sources with at least one author affiliated to UFRJ were retrieved between the last week of January and the first half of February 2023. As the focus was on the institution's scientific output, technological output (e.g., patents) was not retrieved.

The Scopus and Web of Science databases were accessed through the CAPES Periodicals Portal (https://www.periodicos.capes.gov.br). We obtained unrestricted access to the Dimensions interface through its scientometric research support policy (https://www.dimensions.ai/scientometric-research/). Finally, the Lens' academic production retrieval interface (https://www.lens.org/lens/search/scholar/list) required only the creation of a free login to obtain the documents of interest.

All documents were retrieved from the databases' web interfaces. To avoid the inclusion of false positives, the unique identifier 'Affiliation ID' (AF-ID) was used in the Scopus search strategy. Web of Science, on the other hand, has an Affiliation Index that associates variant terms to a canonical institution name. Similarly, the 'Research Organization' field in Dimensions associates all variant terms with a standard institutional name. For Lens, the 'Author Affiliation Name' filter was used with the terms "Federal University of Rio de Janeiro", "UFRJ" and "Federal University of Rio de Janeiro" to select only documents linked to the institution. As the databases restrict the number of publications that can be exported in a single download, we separated the documents into smaller subsets using filters and downloaded them in separate files.

*Data processing*

Once all UFRJ documents had been collected from the four sources, it was necessary to standardize the data due to the discrepancies observed across various fields in the studied databases. The "convert2df" function of the bibliometrix package was used for that end. This function automatically merges all files obtained for a given database into a single table and standardizes multiple fields.

While bibliometrix tools facilitate semi-automated analysis, further standardization was occasionally required to enhance comparison and visualization. One example was the "Document type" field, which is available in all sources, but features a wide variation in the number of categories used in each database to characterize their documents (Dimensions - 5; Scopus - 15; Lens - 18; and WoS - 27). Thus, using the standard classifications of the sources and comparisons between them would not have been feasible. This problem was solved by reducing the number of document types of all the databases to six common categories: (i) Articles; (ii) Books and book chapters; (iii) Event proceedings; (iv) Preprints; (v) Other; (vi) Unidentified. Table 1 presents the category mapping performed to obtain this standardized classification between the different sources.

**Table 1. Merging document categories into a new standardized classification.**

| Original categories | New categories |
|---|---|
| 'Article', 'Journal article', 'Review', 'Article in press', 'Article; Early access', 'Review; Early access', 'Article; Data paper', 'Data paper'; 'Article; Retracted publication', 'Article; Data paper; Early access', 'Reprint' | Articles |
| 'Proceeding', 'Conference proceedings article', 'Conference proceedings', 'Conference paper', 'Conference review', 'Proceedings paper', 'Meeting abstract', 'Article; Proceedings paper' | Proceedings items |
| 'Book', 'Book chapter', 'Chapter', 'Article; Book chapter', 'Review; Book chapter' | Books and book chapters |
| NA | Unidentified |
| 'Preprint' | Preprint |
| 'Editorial material', 'Letter', 'Editorial', 'Note', 'Erratum', 'Book review', 'Correction', 'Short survey', 'Report', 'Other', 'Monograph', 'Biographical-item', 'Dataset', 'Abstract report', 'Discussion', 'Clinical trial', 'Dissertation', 'Reference entry', 'News item', 'Correction, Addition', 'Item about an individual', 'Journal issue', 'Bibliography', 'Editorial material; Book chapter', 'Record review', 'News', 'Art exhibit review', 'Chronology', 'Poetry', 'Retraction' | Other |

Subsequently, a data cleaning process was conducted to eliminate duplicates in the retrieved dataset, which could potentially lead to an overestimation in the results of subsequent analyses. Documents were grouped as duplicates when: (i) all their fields are identical; (ii) they have a duplicate DOI; or (iii) they present identical information for the title, source, author and publication year fields simultaneously.

*Data analysis*

Following the standardization and cleaning of the data, the analysis phase began. We adopted a descriptive statistics methodology that primarily leverages totals and percentages to illustrate and summarize various aspects of the corpus retrieved from each database. The *biblioAnalysis* and *summary* functions, both present in bibliometrix, were used to obtain an initial set of statistics, enabling comparisons among different sources. The entire data analysis and visualization processes were performed using the R language v.4.1.2 (R Core Team, 2023) and the Tidyverse metapackage (Wickham et al., 2019).

For the comparative analysis of document distribution between the sources, we used the R package biblioverlap (Vieira & Leta, 2024). This tool processes two or more bibliographic databases, categorizing documents based on the presence or absence of a unique identifier (such as DOI) and detecting document overlap between datasets when: (i) the identifier is identical for two documents; or (ii) the analysis of Title, Author, Year of Publication and Source fields yields a score that surpasses a specified threshold. The package was used to perform a coverage overlap analysis - mapping documents retrieved from distinct databases to pinpoint those appearing in multiple sources at once - on the entire collection of documents and relevant data subsets corresponding to distinct document types. As database classification discrepancies may lead to the pairing of different document types, we used biblioverlap's *get_all_subset_matches* function to retrieve all paired documents against the subsets of interest before analyzing their overlap.

We used only articles for the comparative analysis of citation-based indicators between databases, as these documents are generally the main source of information in citation analysis (Mingers, Leydesdorff, 2015). This variable was used to compute the following indicators: h (Hirsch, 2005); e (Zhang, 2009); g (Egghe, 2006); $h^c$ (Sidiropoulos et al., 2007) and i10. A review by Garner et al. (2018) defines and presents information about the formulae of all these metrics.

Given that citations increase over time and correlate with the availability of citing documents (Tahamtan, 2016), we also aimed to examine the impact of citation windows and the growth of literature size on these indices. First, we split the articles into five groups according to their publication years: one group for those published before 1983 and four additional groups corresponding to each decade from 1983 through 2022. Then, eight metrics of interest were calculated for the set of documents published in each period, namely: (i) number of articles; (ii) total citations received; (iii) average citations received; (iv) h-index; (v) e-index; (vi) g-index; (vii) $h^c$-index; and (viii) i10-index.

The scripts for data processing and analysis can be found in a public GitHub repository (https://github.com/gavieira/database_coverage_ufrj), which contains thoroughly annotated code that elucidates each step conducted in the process. As Lens allows the redistribution of its data ( https://about.lens.org/policies/#acceptableuse ), the dataset used in this work can be accessed at https://zenodo.org/records/10500802. The datasets downloaded from the other databases are proprietary and, as such, are not available.

## Results

The comparative analyses of UFRJ's scientific output from the selected secondary data sources are organized into three principal sections: (i) an overview of the total number of documents retrieved per database, categorized by publication year and document type; (ii) a document-level coverage analysis of UFRJ's scientific output and the overlap among databases across several data subsets; and (iii) a comparative analysis of citation-based indexes derived from journal articles within each database, examined both collectively and by decade.

*Production retrieved and total publications by year and document type*

We began our data analysis by examining the total number of UFRJ-affiliated documents retrieved by each datasource. Lens indexed a significantly higher number of documents (113,771) compared to the others. Scopus and Dimensions recovered an intermediate number of items (94,472 and 89,327, respectively), while Web of Science (WoS) returned the smallest set (77,143). Altogether, a total of 374,713 documents were recovered.

The next step was to analyze the annual output of UFRJ across these sources. Figure 1 displays the relative frequency of publications by year and data source. In the chart, each bar represents 100% of the documents in a given year, with the proportion of documents from each source differentiated by color. The value within each colored segment indicates the number of publications indexed by the corresponding source in that year. Notably, overlapping documents (i.e., those indexed by multiple sources) are counted in each segment.

Scopus and Dimensions retrieved the most documents published before 1960, although the overall volume of publications during this early period was relatively small. It is also worth noting that UFRJ's scientific output is covered by all four sources only from 1966 onward, when each database indexed the same number of publications (n = 2). From 1967 to 1970, Dimensions and Lens alternated as the source with the highest number of indexed documents.

Between 1971 and the mid-1980s, documents jointly indexed by Web of Science and Scopus constituted a particularly notable portion of the total. From the mid-1980s to around 2000, the number of documents indexed by each database remained fairly consistent. From 2001 onward, Lens indexed a larger share of UFRJ's output, except for 2022, the last year of our analysis, when Dimensions and Lens both retrieved more documents than the other sources. Also in 2022, all databases recorded a decline in total publications compared to the previous year.

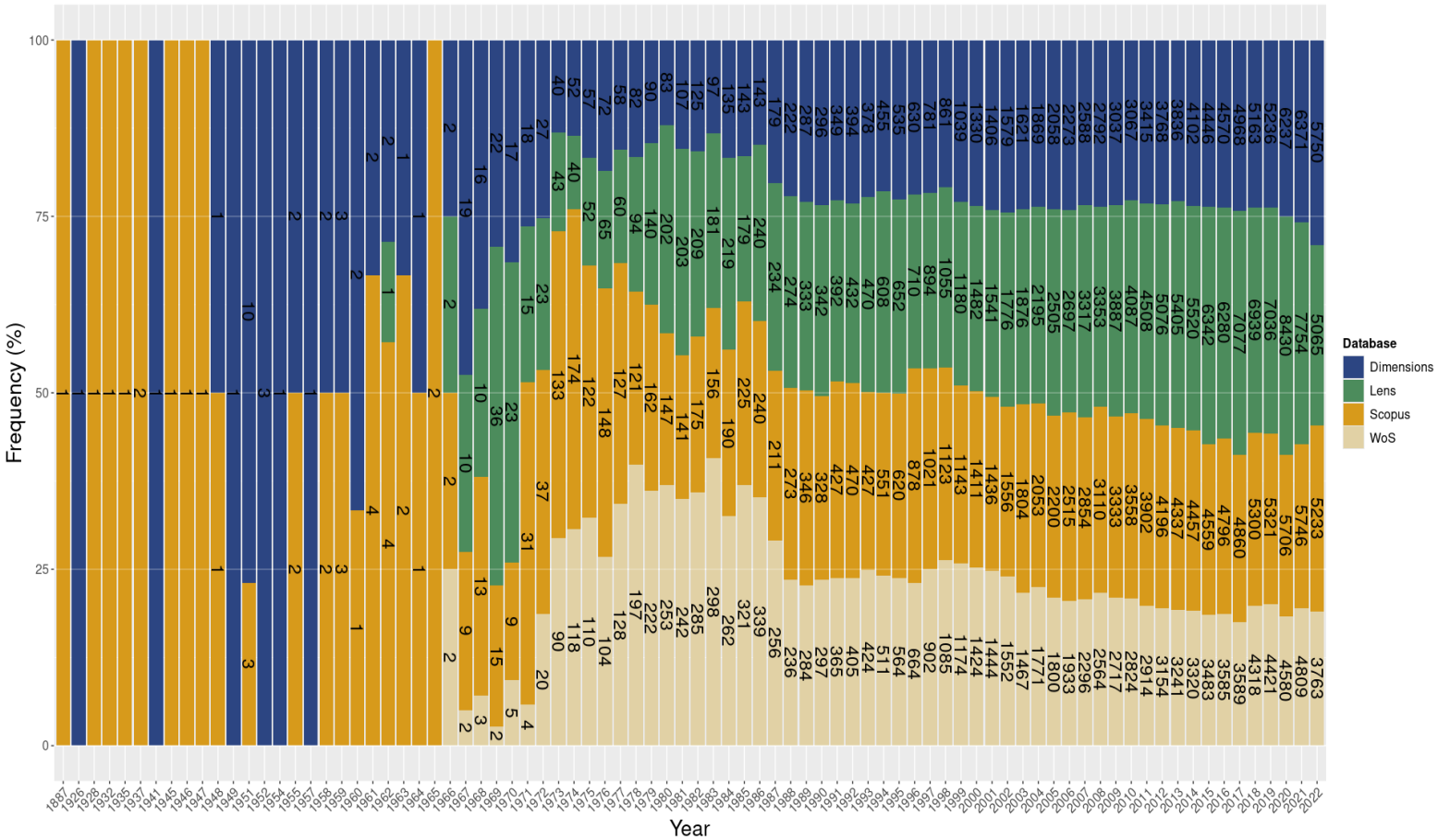


**Figure 1. Relative and total frequency of UFRJ documents by year and data source (1887-2022).**

As previously mentioned, the distinct document type classifications from each database were unified into six standardized categories, enabling a comparative analysis of their occurrence. The total number of documents per category in each secondary source is presented in Figure 2.

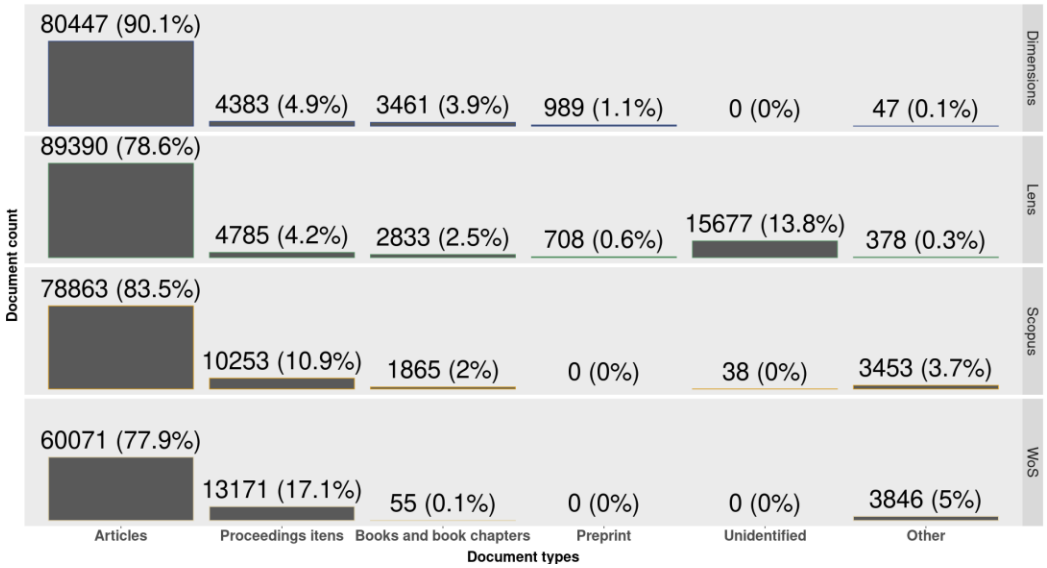


**Figure 2. Total UFRJ documents per standardized document category and data source (1887-2022).**

Most of the documents consist of articles in all databases, ranging between 77.9% and 90.1% of their respective corpus. Regarding the total number of articles, Lens leads with 89,390 indexed items, surpassing Dimensions 80,447 by nearly 9,000, while Scopus follows closely with 78,863 articles, and Web of Science trails with significantly fewer at 60,071. The scenario is quite different when considering the proceedings items. Web of Science indexes the largest number of documents (13,198), followed by Scopus (10,285), whereas Lens and Dimensions contain substantially fewer at 4,787 and 4,400, respectively.

As for books and chapters, Dimensions leads as the most extensive indexer with 3,461 entries, followed by Lens with 2,833, Scopus with 1,865, and Web of Science, where this document type is nearly non-existent, with only 55 indexed records. Preprints are found only in Dimensions (989) and Lens (708), representing a minor fraction of their documents. Unidentified records are predominantly found in Lens, where they are the second most numerous document type, with 15,677 entries, almost 14% of the documents recovered from the database. Scopus also features a small number of unidentified records (38). Finally, the "Other" category is much more prevalent in Web of Science (3,846) and Scopus (3,453) than in Lens (378) and Dimensions (47).

*Coverage analysis by document type*

The *biblioverlap* package was employed to identify the extent of overlap in UFRJ's scientific output across the four datasets. All 374,713 retrieved documents were submitted for analysis. All databases combined would yield 164,366 distinct records, provided overlaps are merged into single entries. Of these, 69,285 documents were found to be exclusive to a single data source, while 95,081 appeared in multiple sources. Among the overlapping documents, 92,314 were matched via DOI, whereas the remaining 2,767 were identified through comparative analysis of other bibliographic fields - specifically, title, publication year, first author's name, and journal title.

The results of this coverage analysis were also used to generate Venn diagrams at three distinct aggregation levels: (i) the complete dataset; (ii) subsets based on document type; and (iii) the subset containing the most cited articles, defined here as those that belong to the h-core of each database. The diagrams obtained are shown in Figure 3 and illustrate the document-level coverage overlap in all databases examined. The analysis includes the full set of retrieved documents (3A), records classified as 'articles' (3B), h-core articles for each source (3C), items categorized as 'conference items' (3D), 'books and chapters' (3E), and 'other' types (3F). Each intersection displays the number of documents it contains, followed by the percentage this represents relative to the total number of distinct records analyzed (164,366). The shading of each intersection reflects the number of documents it contains: darker shades correspond to higher values relative to other intersections, while lighter shades to lower values.

The analysis of the full dataset (Fig. 3A) reveals that over a quarter of the documents (45,242) are present across all four databases, whereas those shared between two or three databases are significantly fewer, not exceeding 5% of the total distinct documents. The only exceptions are the document sets found concurrently in Lens, Scopus, and Dimensions (14,021 - 9%) and those in Lens and Dimensions (13,651 - 8%). Regarding documents that occur exclusively in one database, Lens leads with 31,272 records (19%), surpassing even the combined counts from Scopus (16,110 - 10%) and Web of Science (14,819 - 9%), whereas Dimensions contains the fewest exclusive items at 7,084 (4%).

For articles (Fig. 3B), there is a considerable decrease in the proportion of exclusive documents from Scopus, Web of Science, and, above all, Lens. Also, the fraction of articles that occur simultaneously in all sources is bigger (from 28% when analyzing all documents to 36% when analyzing only articles). For h-core articles (Fig. 3C), the percentage of items shared by all datasets is even higher (49%).

Regarding conference items (Fig. 3D), a substantial proportion (11%) is found across all four databases. The only other intersection with a notable share is the one comprising Dimensions, Lens, and Scopus (12%). Beyond these two cases, the presence of conference items in multiple sources is relatively limited, with no other intersection exceeding 5% of the total in this subset. Exclusivity is also prominent in this category, particularly in Scopus (14%) and Web of Science (34%), which hold the largest shares of conference items not indexed by other databases.

For books and chapters (Fig. 3E), Scopus and Dimensions stand out with relatively high proportions of exclusive content - 17% and 13%, respectively -, followed by

Lens at 9%. Conversely, only a very small portion of these documents (1%) is shared across all four databases. The intersection encompassing Lens, Scopus, and Dimensions accounts for the largest share within this category, representing 33% of the subset.

Finally, the 2,107 documents classified under the 'Other' category are simultaneously indexed by all databases (Fig. 3F) - a figure notably higher than the totals reported in the original classifications provided by Lens and Dimensions.

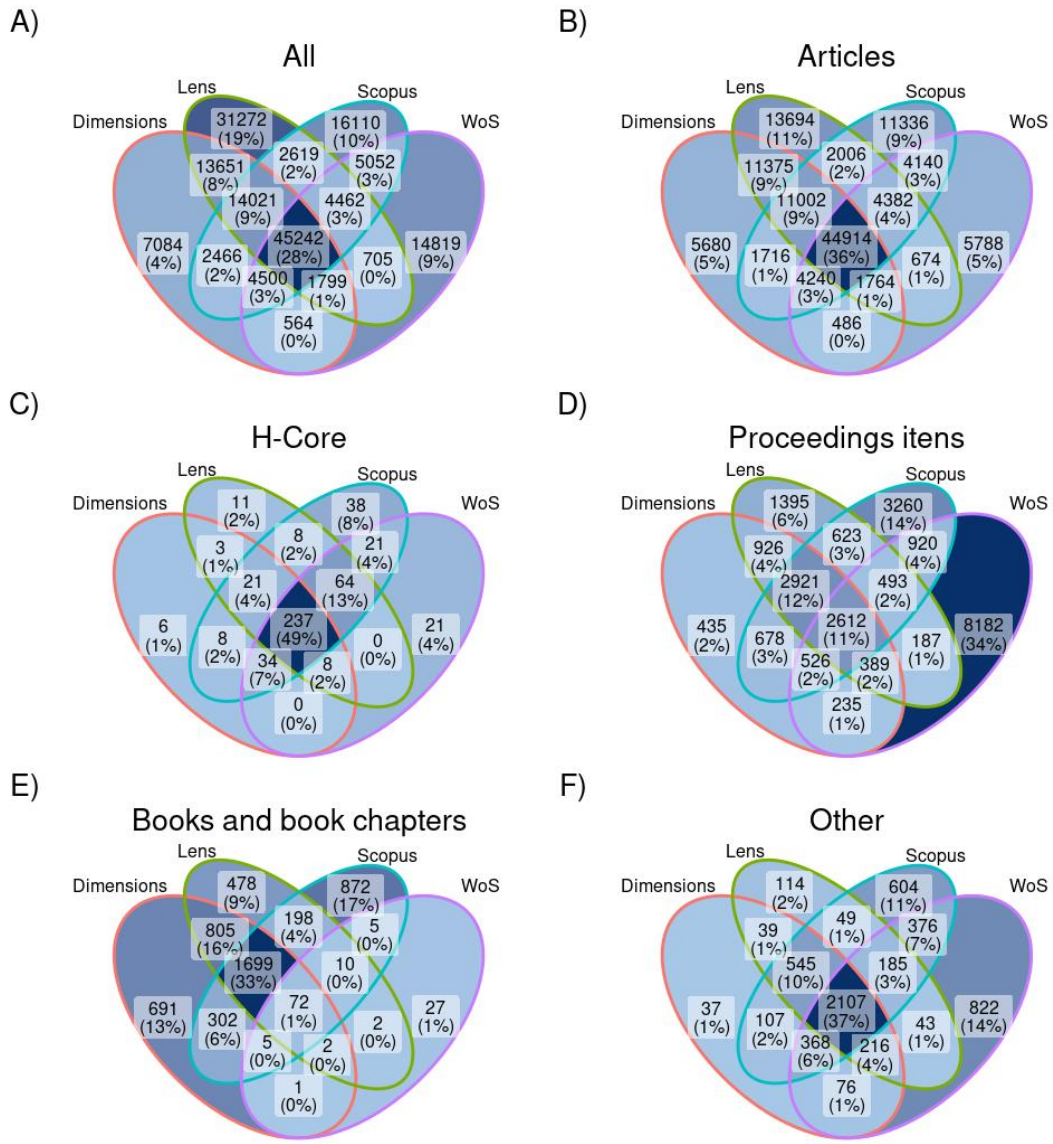


**Figure 3. Venn diagrams representing UFRJ's scientific production overlap between the bibliographic databases for multiple data subsets (1887-2022).**

*Citation-based indicators in each source*

In addition to the comparative analysis on the entire corpus of UFRJ publications, five citation-based indices (h, e, g, $h^c$ and i10) were calculated for the set of articles recovered by each database. Table 2 presents these indices along with the total counts of articles and citations.

In general, a higher total article count does not necessarily translate into higher indicator values: if that were the case, Lens would show the highest values and Web of Science the lowest. However, for most indicators, we found that Scopus yields the highest values, followed by Lens, Web of Science, and finally Dimensions. A particularly clear example of this can be seen when comparing Dimensions and Web of Science: although Dimensions has approximately 20,000 more articles and 112,000 more citations than Web of Science, it shows slightly lower values across all indices, except for the i10 index.

**Table 2. Total number of items, citations received and visibility indicators calculated based on UFRJ articles retrieved from each data source (1887-2022).**

| Source | No. of articles | Total citations | Index h | Index e | Index g | Index $h^c$ | i10 Index |
|---|---|---|---|---|---|---|---|
| Lens | 89.390 | 1.539.910 | 307 | 322,18 | 500 | 159 | 33.310 |
| Dimensions | 80.447 | 1.376.751 | 279 | 278,05 | 442 | 145 | 32.240 |
| Scopus | 78.863 | 1.642.724 | 316 | 353,89 | 533 | 164 | 35.984 |
| WoS | 60.071 | 1.264.978 | 281 | 304,86 | 466 | 147 | 28.636 |

Citation-based indexes and other metrics of interest were also computed by decade (Figure 4). Most of the indices follow a specific trend, regardless of the source: the set of oldest articles (pre-1983) displays the lowest indices, which increase subtly in the period between 1983 and 1992. These indices are significantly higher for articles from the next decade (1993-2002) and continue to grow at the same rate in the period between 2003 and 2012. Then it falls slightly for articles published between 2013 and 2022. Some indicators, such as the i10 index, have diverged from this trend in the last decade analyzed by maintaining their value, while the $h^c$ index has exhibited increments during the same period.

Though all databases follow this pattern, there were still differences in their results. For instance, Scopus performs slightly better than the other databases in most metrics, especially for the last two decades analysed. The exceptions to that are the total number of indexed documents, which are higher in Lens and Dimensions, and mean citations, where WoS outperforms all the other databases.

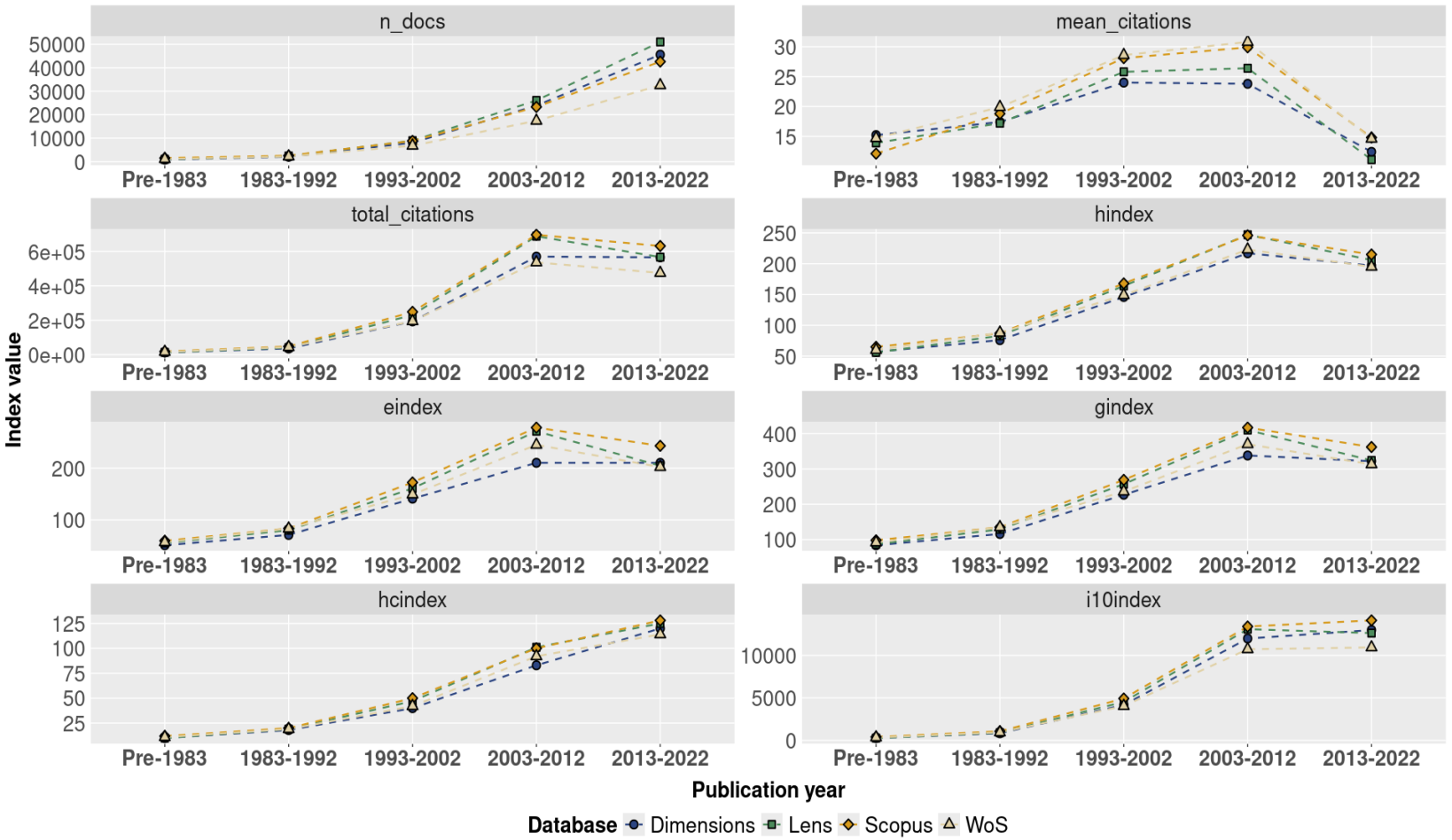


**Figure 4. Total number of items, total and average citations received, and visibility indicators calculated for the entire output of UFRJ articles in each data source, grouped by publication year windows (1887-2022).**

## Discussion

The **total number of documents** retrieved generally correlates with the total indexed documents in each source (Gusenbauer, 2022), indicating that a higher volume of documents increases the likelihood of retrieving a substantial number of relevant documents. However, Dimensions is an exception to this rule. Previous research shows that Dimensions offers significantly broader coverage of publications (Visser et al., 2021) and journals (Singh et al., 2021) compared to Scopus and Web of Science. Notwithstanding, Dimensions retrieved 5,000 less UFRJ-associated documents than Scopus.

This unexpected result may be linked to the substantial proportion of Dimensions documents that are either unaffiliated with any country or institution or lack complete affiliation data (Guerrero-Bote et al., 2021). As this work depends on the quality of the affiliation for data retrieval, it makes sense that Dimensions returned fewer results than Scopus, even though it indexes more documents.

Regarding the distribution of **documents by year**, Scopus and Dimensions retrieve the highest number of records published before 1960. While these documents represent only a small fraction of each database's total corpus, they may be valuable for historiographic studies (Thakuria et al., 2024; Ullah et al., 2023) and related fields. A particularly noteworthy case is a publication from 1887 titled *"The Genesis of the Diamond"*, retrieved from Scopus (Derby, 1887). This stands out because UFRJ was founded in 1920, so any publication predating that year could suggest a metadata error. However, upon examining the document, it was found to be a letter published in *Science*, linked to the National Museum - an institution established in 1818 and later incorporated into UFRJ. Thus, the indexing by Scopus is valid, demonstrating a level of curation quality in this database.

The Lens database indexes a larger proportion of documents from the 21st century compared to the other three sources. This percentage has grown steadily in more recent years until it dropped sharply in 2022, when both Dimensions and Scopus retrieved more documents than Lens. This likely reflects the timing of data collection - early 2023 - when none of the databases had fully indexed the previous year's publications. This is supported by the fact that all databases showed a decrease in the number of documents for 2022 compared to 2021, with the decline being especially marked for Lens. Additionally, Lens was the only database to show a drop in the number of documents published in 2021 relative to 2020. This suggests that Lens may have a slower indexing process, which could be a critical factor for assessments focused on recent literature. However, updated data and more detailed analyses would be necessary to confirm this.

It is also worth noting that, despite its known issue of incomplete affiliation data, Dimensions has consistently indexed more UFRJ-affiliated documents than Scopus since 2020. Two potential hypotheses could explain this: (i) improvements in Dimensions' indexing practices, especially for recent publications and/or (ii) an increase in the volume of content indexed by Dimensions in comparison to Scopus, resulting in a greater number of documents even if there were no improvements to its indexing methodology. Again, further analysis would be required to investigate these claims.

The majority of **documents types** in all four databases are classified as journal articles. However, there is a gap of more than 9,000 articles between the database with the highest article count, Lens, and the second-highest, Scopus. This difference is even more pronounced when compared to the Web of Science (WoS), which has roughly a third as many articles as Lens. While a larger volume of indexed articles can be an attractive characteristic when selecting a database for institutional evaluation, it should not be the sole criterion. Other factors, such as the disciplinary focus of the evaluation and the relevance and prevalence of specific document types, must also be taken into account.

Proceedings items offer a compelling example of this point. Scopus and WoS significantly outperform the newer sources in terms of total indexed records - a pattern that contrasts sharply with what is observed for journal articles. This is especially relevant for evaluating disciplines where research dissemination is more dynamic, such as computer science and related subfields like human-computer interaction, where proceedings are a primary channel for communicating new findings (Freyne et al., 2010; Meho & Rogers, 2008). In such fields, a high count of journal articles may not compensate for a poor representation of proceedings.

In contrast, Dimensions and Lens index more documents classified as books or book chapters. WoS, on the other hand, includes very few of these document types. This limitation may impact its effectiveness in evaluating disciplines where books remain a key vehicle for scholarly communication, which is generally the case for the social sciences and humanities (Kousha & Thelwall, 2015; Bornmann et al., 2016; Toledo, 2020).

As for preprints, it is worth highlighting their growing importance as a means of accelerating the dissemination of research results. While not peer-reviewed, preprints have been especially valuable in contexts that require rapid knowledge sharing, such as during the COVID-19 pandemic (Fraser et al., 2021). However, only Dimensions and Lens include preprints, and even then, they account for only a small fraction of each database's contents.

Unidentified documents were detected in both Lens and Scopus, though their number is negligible in the latter. In contrast, Lens contains a considerable proportion of documents (13.8%) that lack an assigned document type, indicating a potential shortcoming in its classification system..

Documents classified as "other" are abundant in WoS and Scopus, but scarce in Lens and, especially, Dimensions. Differences in how databases classify their content may explain this variation, since this category typically aggregates a large number of widely diverse document types that do not have a direct match in all the sources. Likely, some documents labeled as "Other" in WoS and Scopus are assigned to other categories (e.g., "Articles" or "Conference Items") in Dimensions and Lens. A clearer understanding of these discrepancies, however, requires a closer examination through the subsequent coverage analysis.

Numerous studies have highlighted the considerable variation in coverage overlap of scientific output across different bibliographic databases (Gusenbauer, 2019; Huang et al., 2020; Martín-Martín et al., 2021; Visser et al., 2021). Given this methodology's popularity and usefulness, we have employed a **coverage overlap**

**analysis** to better characterize UFRJ's scientific output in the four databases analyzed.

As expected, most document matches were made via DOI - a result consistent with previous studies, such as Visser et al. (2021), where 80% or more of the matched documents across sources were linked through DOI-based filters. The remaining matches in those studies were based on combinations of bibliographic metadata like first author's surname and year of publication. In our case, more than two thousand documents without a DOI were still successfully matched using other bibliographic fields. This underscores one of the key advantages of the biblioverlap package, which prioritizes matching via a unique identifier (like DOI) but falls back on a scoring mechanism based on multiple metadata fields when such identifiers are absent (Vieira & Leta, 2024). This approach minimizes matching data loss and is particularly valuable in a multi-disciplinary analysis that includes diverse document types, since DOI assignment practices can vary widely across fields (Gorraiz et al., 2016).

Despite being the database with the largest dataset, Lens presented an unexpectedly high number of unique documents (31,272). Upon manual inspection, we have found that 15,492 of these belonged to entries lacking document type classification. It's worth emphasizing two points here: (i) these documents were generally not associated with DOIs, and (ii) to improve computational efficiency, the biblioverlap algorithm assumes that if a document has a unique identifier (such as a DOI), it will be present in all datasets being analyzed. Consequently, some of the documents deemed unique to Lens may have DOIs in other databases, DOIs that Lens failed to capture, potentially inflating the count of supposedly exclusive documents.

The observed decrease in the number of exclusive documents for the "article" type is largely attributable to the exclusion of document types containing high proportions of unique entries, namely the unidentified documents in Lens and the conference items in Scopus and WoS. In fact, around one-third of journal articles are retrievable from any of the databases, and the number of truly exclusive articles is relatively low. This indicates that, for evaluations centered on journal articles, database coverage alone may not be a distinguishing factor. Other aspects - such as metadata quality and available bibliographic fields - may be more relevant when selecting a source.

We also examined highly cited documents in each database's h-core. Notably, half of these (237) were retrieved by all four sources. This aligns with Visser et al. (2021), who showed that more highly cited documents tend to appear across multiple databases. Since citations are influenced by journal prestige (Martin & Irvine, 1983; Bornmann et al., 2012), and highly cited journals are often prioritized for indexing (Garfield, 1999), it's expected that these publications will appear in multiple sources. Proceeding items show a markedly different pattern: a low overlap across sources and a high number of documents exclusive to one database. WoS retrieves the largest share of these documents (around 57%), and combining WoS and Scopus increases this to approximately 89% of all UFRJ-affiliated conference items. This supports the use of both databases in evaluations of disciplines where proceedings are a key publication venue.

About one-third of books and book chapters were found simultaneously in Scopus, Dimensions, and Lens - a direct consequence of the near absence of this document type in WoS. The limited increase in the total number of WoS records after accounting for matches with other document types supports the conclusion that this is a real coverage gap rather than a classification issue. Thus, WoS may not be suitable for evaluation processes focused on areas where books are particularly relevant for scientific communication.

The "Other" document type presents another intriguing case. In Dimensions and Lens, the number of documents matched to the entries categorized as "Other" by the remaining databases was far greater than the number of items those sources originally classified as such. Manual inspection revealed that the majority of these documents were classified as "Articles" in Dimensions (97.8%) and Lens (88.2%), suggesting classification errors. While further investigation would be needed to determine definitively which databases are misclassifying documents, Dimensions and Lens, being relatively recent and drawing heavily from open data aggregators like PubMed and Crossref (Herzog et al., 2020; Cambia, 2024a), are more likely to be the sources of these inconsistencies. Metadata quality from Scopus and WoS is generally regarded as more reliable (Guerrero-Bote et al., 2021; Delgado-Quirós & Ortega, 2024).

Such classification issues have serious implications for scientometric analyses. As previously discussed, the impact and relevance of document types vary greatly by discipline. Misclassifications can skew evaluations or lead to erroneous conclusions. Therefore, while total document count is an important metric when choosing a data source, harder-to-measure qualities - like metadata accuracy and classification reliability - are just as critical, if not more so.

The h-index combines publication and citation counts into a single metric (Hirsch, 2005), favoring documents with higher citation volumes. This attribute is shared by its derivatives, such as the e-, g-, and hc-indexes. As a result, smaller databases may yield higher values for **citation-based indices** values than larger ones, depending on how well they capture highly cited publications.

Our findings reflect this pattern. WoS outperformed Dimensions, and Scopus outperformed Lens, despite the latter two having broader overall coverage. We suggest two main aspects that could explain this outcome: (i) the more performant databases may include highly cited documents absent from other sources; or (ii) they may have more efficient citation-linking mechanisms. The first seems less likely, as our overlap analysis showed that approximately half of the h-core documents are shared across all platforms.

The second aspect is more plausible. Issues with metadata precision appear to hinder accurate citation tracking in the newer sources. For example, Lens has a substantial number of uncategorized documents, suggesting a lack of granularity in its curation process. Similarly, despite its large document base, Dimensions retrieved fewer publications, likely due to deficiencies in the 'affiliation' field. Visser et al. (2021) corroborates this view by reporting that, while highly cited articles tend to be present in all major databases, WoS and Scopus demonstrate superior citation-linking capabilities. Our results are consistent with these findings.

An exception to this pattern is the i10-index, which counts the number of publications with at least ten citations. Unlike other indices, it is not increasingly difficult to raise its value over time. Because it uses a fixed, relatively low citation threshold, it is also less sensitive to errors in citation linking. Notably, this was the only index where Dimensions outperformed WoS. This underscores the importance of carefully selecting citation metrics, as different indices may produce varying outcomes depending on the data source and characteristics of the dataset.

When analyzing UFRJ's scientific output by decade, we found that citation-based indices are shaped by both the volume and age of publications. Early periods had few indexed articles and lower index values, even though these documents had more time to accumulate citations. Later decades featured both an increase in publication volume and more extensive citation windows, which corresponded to higher index values. In the most recent decade, although publication volume kept growing, the indices plateaued or declined - likely a result of limited time for newer publications to accrue citations.

Interestingly, the i10 and $h^c$ indexes did not decline in the most recent decade. The i10-index's resilience likely reflects its modest citation threshold, though Lens saw a drop that may be linked to slower indexing of recent publications. The hc-index, which gives greater weight to recent citations, actually increased, as expected.

Together, these results demonstrate that citation-based index values are not only influenced by the selected database but are also highly dependent on the metric chosen. Scopus consistently delivered higher index values, likely due to its balanced combination of broad coverage and efficient citation linking. WoS, while similarly strong in citation linking, indexes a more selective subset of publications. This results in higher average citation values but not necessarily higher index values. By contrast, Dimensions and Lens reported lower mean citation values, pointing to either less effective citation tracking, the inclusion of more poorly cited documents, or both.

## Conclusion

Although this study offers important insights into how crucial database selection may be to institutional research evaluation, it has several limitations. First, it is based on a single case study and considers only two variables: document counts and citation counts. Furthermore, it does not split the production by discipline and evaluates only four bibliographic databases.

The analyses were conducted using the complete scientific output of one university - UFRJ. We make no claims that these results are generalizable to other institutions, and we recognize that similar analyses may yield different results elsewhere. For transparency and reproducibility, the datasets used (where legally permissible) and the analysis code have been made publicly available on Zenodo and GitHub. We hope this facilitates the application of our analytical framework to other institutions, encouraging replication, validation, or expansion of our findings while addressing different dimensions of institutional evaluation. In fact, we are currently conducting a follow-up study using a similar methodology to examine both high- and low-ranked institutions, aiming to determine whether characteristics such as publication overlap correlate with institutional reputation.

An additional limitation lies in the narrow focus on two variables, document and citation counts, which excludes other relevant dimensions of research output. These include collaboration patterns (both national and international) and adherence to open access publishing models, both of which are available in the databases analyzed. Future studies would benefit from incorporating these additional dimensions alongside citation-based indices. Such multifaceted analyses could support the development of custom indicators or institutional rankings, as seen in the work of Huang et al. (2020).

Another constraint is the absence of field-level classification in our analysis. Categorizing scientific output by discipline is a valuable addition to any bibliometric study, allowing for finer-grained comparisons, particularly when interpreting citation-based indicators. Including such classification in future work - whether by mapping categories across databases (Singh et al., 2021) or through publication-level classification based on content (Rivest et al., 2021; Pech et al., 2022) - would enhance analytical depth.

Disciplinary classification would also allow for normalization of citation-based indicators by field (Waltman & van Eck, 2013) and enable more appropriate comparisons across disciplines. Moreover, it would support the exploration of domain-specific patterns, such as citation half-life (Burton & Kebler, 1960), which varies significantly across research areas. With publications sorted by discipline, it becomes possible to assess how citation window length affects the evaluation of different fields across various sources. Ultimately, this would improve our understanding of how database and indicator choices may affect evaluations not only at the institutional level but also within specific disciplines.

Lastly, the decision to analyze four databases was a methodological choice driven by the scope of the study. We opted to examine the full scholarly output of a large institution from its founding up to the year before data collection. Expanding the analysis to include more sources would have significantly increased the data volume and required much more time for retrieval and processing. However, we intend to include additional databases in future research. OpenAlex - a relatively recent, open-access, multidisciplinary source - stands out as a promising candidate due to its publicly available API and user-friendly interface (Priem et al., 2022).

The primary objective of this study was to examine how the choice of bibliographic database can affect both the set of retrieved publications and the calculation of citation-based indicators, as well as to discuss the broader implications for evaluating an institution's scientific output. To achieve this, we conducted a case study using the complete scholarly production of the Federal University of Rio de Janeiro (UFRJ), as retrieved from four multidisciplinary databases: Web of Science (WoS), Scopus, Dimensions, and Lens.

Our analysis of UFRJ's production across these databases revealed several findings that support the notion that results vary substantially depending on the source, particularly in terms of total document count, document type coverage, and citation-based metrics. One key takeaway is that a larger database does not necessarily guarantee higher retrieval of relevant documents or better citation metrics, as factors such as metadata quality are equally, if not more, important than the number of indexed items.

For instance, although Dimensions is considerably larger than Scopus in terms of overall indexed content (Visser et al., 2021), it retrieved fewer documents in this case study. This is likely due to limitations in the quality of metadata, especially within the 'affiliation' field (Guerrero-Bote et al., 2021), which was used as the main retrieval criterion. Similarly, Lens displayed signs of metadata quality issues, such as a high proportion of records lacking classification by document type. When analyzing citation-based indices, we found that databases with higher retrieval counts did not necessarily yield higher index values: Lens, for example, was outperformed by Scopus, while Dimensions was outperformed by WoS. These findings likely reflect differences in the efficiency of citation link identification across databases.

In summary, although newer data sources, like Dimensions and Lens, tend to be more comprehensive in terms of document indexing, they also show signs of lower metadata quality when compared to more established sources like WoS and Scopus. Therefore, selecting a database solely based on its volume of indexed content may be inadvisable. Because scientific output evaluations can be significantly influenced by both metadata quality and coverage of specific document types, neglecting these characteristics may lead to inaccurate assessments. Similarly, selecting appropriate citation-based indices is essential to avoid distortions caused by, for instance, a few highly cited publications.

Our coverage analysis also showed that a significant proportion of documents - especially journal articles and highly cited papers - appear in all four databases. However, there was wide variation in document type coverage, which is relevant given that different disciplines often rely on different formats for scholarly communication. These results further underscore the inadequacy of database selection based on document count alone.

In the context of our dataset, Scopus emerged as the most suitable database in terms of both document retrieval and citation-based indicators. This finding aligns with the methodologies of prominent university rankings such as QS and THE, both of which use Scopus as the underlying data source for evaluating institutional output. Additionally, Scopus included a large number of items classified as "Books or Chapters," achieving comparable coverage of this document type to that of Dimensions and Lens. Considering the lower metadata quality observed in the newer sources (Guerrero-Bote et al., 2021; Visser et al., 2021), Scopus appears particularly well-suited for evaluating fields in which books constitute a key channel for scholarly communication.

Nevertheless, this does not imply that other databases should be disregarded. Dimensions, for instance, integrates data on publications, altmetrics, clinical trials, patents, funding, and institutional policies, which are interlinked through citations and other types of connections. According to Herzog et al. (2020), these relationships enable a holistic analysis of the scientific production cycle: from initial research funding to publication, technological application, and influence on policy development. The authors also note that the developers of Dimensions are aware of the limitations associated with this database and are actively working with publishers and other partners to enhance its content quality and coverage. The same can likely be said for Lens. Given that both databases were launched in the late 2010s,

significant improvements in their performance and coverage can be expected over time.

Although WoS indexed fewer documents than Scopus, it demonstrated a more granular classification of document types. It also outperformed Dimensions, despite the latter's larger size, which suggests that WoS is highly efficient in establishing citation links among its records. Moreover, WoS's rich metadata and extensive collection of "Conference items" indicate that it is a valuable resource for evaluating disciplines where such formats are a key channel of scientific communication.

We hope that the findings presented here raise awareness among researchers, evaluators, and policymakers regarding how both database and metric selection can significantly affect institutional assessments. Recognizing these effects is a critical step toward promoting higher standards in research evaluation and ensuring that methodologies are appropriately tailored to the specific characteristics and needs of each evaluation context.

**Acknowledgements**

We would like to thank CNPq for their continued financial support.

## References

Archambault, É., Campbell, D., Gingras, Y., & Larivière, V. (2009). Comparing bibliometric statistics obtained from the Web of Science and Scopus. Journal of the American Society for Information Science and Technology, 60(7), 1320–1326. https://doi.org/10.1002/asi.21062

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics, 11(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

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

Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. Quantitative Science Studies, 1(1), 363–376. https://doi.org/10.1162/qss_a_00018

Bornmann, L. (2018). Field classification of publications in Dimensions: A first case study testing its reliability and validity. Scientometrics, 117(1), 637–640. https://doi.org/10.1007/s11192-018-2855-y

Bornmann, L., Haunschild, R., & Mutz, R. (2021). Growth rates of modern science: A latent piecewise growth curve approach to model publication numbers from established and new literature databases. Humanities and Social Sciences Communications, 8(1), Article 1. https://doi.org/10.1057/s41599-021-00903-w

Bornmann, L., Schier, H., Marx, W., & Daniel, H.-D. (2012). What factors determine citation counts of publications in chemistry besides their quality? Journal of Informetrics, 6(1), 11–18. https://doi.org/10.1016/j.joi.2011.08.004

Bornmann, L., Thor, A., Marx, W., & Schier, H. (2016). The application of bibliometrics to research evaluation in the humanities and social sciences: An exploratory study using normalized Google Scholar data for the publications of a research institute. Journal of the Association for Information Science and Technology, 67(11), 2778–2789. https://doi.org/10.1002/asi.23627

Burton, R. E., & Kebler, R. W. (1960). The "half-life" of some scientific and technical literatures. American Documentation, 11(1), 18–22. https://doi.org/10.1002/asi.5090110105

Carlsson, H. (2009). Allocation of Research Funds Using Bibliometric Indicators – Asset and Challenge to Swedish Higher Education Sector. Infotrend, 64(4), 82–88.

Chadegani, A. A., Salehi, H., Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., & Ebrahim, N. A. (2013). A Comparison between Two Main Academic Literature Collections: Web of Science and Scopus Databases. Asian Social Science, 9(5), Article 5. https://doi.org/10.5539/ass.v9n5p18

de Oliveira, T. M., & Amaral, L. (2017). Políticas Públicas em Ciência e Tecnologia no Brasil: Desafios e propostas para utilização de indicadores na avaliação. In Bibliometria e Cientometria no Brasil: Infraestrutura para avaliação da pesquisa científica na Era do Big Data (pp. 157–184). ECA/USP.

Delgado-Quirós, L., Aguillo, I. F., Martín-Martín, A., López-Cózar, E. D., Orduña-Malea, E., & Ortega, J. L. (2023). Why are these publications missing? Uncovering the reasons behind the exclusion of documents in free-access scholarly databases. Journal of the Association for Information Science and Technology, asi.24839. https://doi.org/10.1002/asi.24839

Delgado-Quirós, L., & Ortega, J. L. (2024). Completeness degree of publication metadata in eight free-access scholarly databases. Quantitative Science Studies, 1–19. https://doi.org/10.1162/qss_a_00286

Derby, O. A. (1887). The Genesis of the Diamond. Science, ns-9(207), 57–58. https://doi.org/10.1126/science.ns-9.207.57

Egghe, L. (2006). Theory and practise of the g-index. Scientometrics, 69(1), 131–152. https://doi.org/10.1007/s11192-006-0144-7

Franceschet, M. (2009). A comparison of bibliometric indicators for computer science scholars and journals on Web of Science and Google Scholar. Scientometrics, 83(1), 243–258. https://doi.org/10.1007/s11192-009-0021-2

Fraser, N., Brierley, L., Dey, G., Polka, J. K., Pálfy, M., Nanni, F., & Coates, J. A. (2021). The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. PLOS Biology, 19(4), e3000959. https://doi.org/10.1371/journal.pbio.3000959

Freyne, J., Coyle, L., Smyth, B., & Cunningham, P. (2010). Relative status of journal and conference publications in computer science. Communications of the ACM, 53(11), 124–132. https://doi.org/10.1145/1839676.1839701

Garfield, E. (1999). Journal impact factor: A brief review. CMAJ, 161(8), 979–980.

Garner, R. M., Hirsch, J. A., Albuquerque, F. C., & Fargen, K. M. (2018). Bibliometric indices: Defining academic productivity and citation rates of researchers, departments and journals. Journal of Neurointerventional Surgery, 10(2), 102–106. https://doi.org/10.1136/neurintsurg-2017-013265

Gorraiz, J., Melero-Fuentes, D., Gumpenberger, C., & Valderrama-Zurián, J.-C. (2016). Availability of digital object identifiers (DOIs) in Web of Science and Scopus. Journal of Informetrics, 10(1), 98–109. https://doi.org/10.1016/j.joi.2015.11.008

Grindlay, D. J. C., Brennan, M. L., & Dean, R. S. (2012). Searching the Veterinary Literature: A Comparison of the Coverage of Veterinary Journals by Nine Bibliographic Databases. Journal of Veterinary Medical Education, 39(4), 404–412. https://doi.org/10.3138/jvme.1111.109R

Guerrero-Bote, V. P., Chinchilla-Rodríguez, Z., Mendoza, A., & de Moya-Anegón, F. (2021). Comparative Analysis of the Bibliographic Data Sources Dimensions and

Scopus: An Approach at the Country and Institutional Levels. Frontiers in Research Metrics and Analytics, 5. https://www.frontiersin.org/article/10.3389/frma.2020.593494
Guerrero-Bote, V. P., & Moya-Anegón, F. (2012). A further step forward in measuring journals' scientific prestige: The SJR2 indicator. Journal of Informetrics, 6(4), 674–688. https://doi.org/10.1016/j.joi.2012.07.001
Gusenbauer, M. (2019). Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. Scientometrics, 118(1), 177–214. https://doi.org/10.1007/s11192-018-2958-5
Gusenbauer, M. (2022). Search where you will find most: Comparing the disciplinary coverage of 56 bibliographic databases. Scientometrics, 127(5), 2683–2745. https://doi.org/10.1007/s11192-022-04289-7
Herzog, C., Hook, D., & Konkiel, S. (2020). Dimensions: Bringing down barriers between scientometricians and data. Quantitative Science Studies, 1(1), 387–395. https://doi.org/10.1162/qss_a_00020
Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America, 102(46), 16569–16572. https://doi.org/10.1073/pnas.0507655102
Huang, C.-K. (Karl), Neylon, C., Brookes-Kenworthy, C., Hosking, R., Montgomery, L., Wilson, K., & Ozaygen, A. (2020). Comparison of bibliographic data sources: Implications for the robustness of university rankings. Quantitative Science Studies, 1(2), 445–478. https://doi.org/10.1162/qss_a_00031
Johnes, J. (2018). University rankings: What do they really show? Scientometrics, 115(1), 585–606. https://doi.org/10.1007/s11192-018-2666-1
Jones, T., Huggett, S., & Kamalski, J. (2011). Finding a Way Through the Scientific Literature: Indexes and Measures. World Neurosurgery, 76(1), 36–38. https://doi.org/10.1016/j.wneu.2011.01.015
Kousha, K., & Thelwall, M. (2015). Web indicators for research evaluation. Part 3: Books and non standard outputs. Profesional de La Información, 24(6), Article 6. https://doi.org/10.3145/epi.2015.nov.04
Liang, Z., Mao, J., Lu, K., & Li, G. (2021). Finding citations for PubMed: A large-scale comparison between five freely available bibliographic data sources. Scientometrics, 126(12), 9519–9542. https://doi.org/10.1007/s11192-021-04191-8
Martin, B. R., & Irvine, J. (1983). Assessing basic research: Some partial indicators of scientific progress in radio astronomy. Research Policy, 12(2), 61–90. https://doi.org/10.1016/0048-7333(83)90005-7
Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado López-Cózar, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. Scientometrics, 126(1), 871–906. https://doi.org/10.1007/s11192-020-03690-4
Meho, L. I., & Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: A comparison of Scopus and Web of Science. Journal of the American Society for Information Science and Technology, 59(11), 1711–1726. https://doi.org/10.1002/asi.20874
Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. European Journal of Operational Research, 246(1), 1–19. https://doi.org/10.1016/j.ejor.2015.04.002
Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. Scientometrics, 106(1), 213–228. https://doi.org/10.1007/s11192-015-1765-5

Pech, G., Delgado, C., & Sorella, S. P. (2022). Classifying papers into subfields using Abstracts, Titles, Keywords and KeyWords Plus through pattern detection and optimization procedures: An application in Physics. Journal of the Association for Information Science and Technology, 73(11), 1513–1528. https://doi.org/10.1002/asi.24655

Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts (No. arXiv:2205.01833). arXiv. https://doi.org/10.48550/arXiv.2205.01833

QS. (2024, July 15). South America QS World University Rankings 2024. Top Universities. https://www.topuniversities.com/latin-america-south-america-rankings

Rivest, M., Vignola-Gagné, E., & Archambault, É. (2021). Article-level classification of scientific publications: A comparison of deep learning, direct citation and bibliographic coupling. PLOS ONE, 16(5), e0251493. https://doi.org/10.1371/journal.pone.0251493

Schneider, J. W. (2009). An Outline of the Bibliometric Indicator Used for Performance-Based Funding of Research Institutions in Norway. European Political Science, 8(3), 364–378. https://doi.org/10.1057/eps.2009.19

Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2007). Generalized Hirsch h-index for disclosing latent facts in citation networks. Scientometrics, 72(2), 253–280. https://doi.org/10.1007/s11192-007-1722-z

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(6), 5113–5142. https://doi.org/10.1007/s11192-021-03948-5

Tahamtan, I., Safipour Afshar, A., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. Scientometrics, 107(3), 1195–1225. https://doi.org/10.1007/s11192-016-1889-2

Thakuria, A., Chakraborty, I., & Deka, D. (2024). A bibliometric review on serendipity literature available in Web of Science database using HistCite and Biblioshiny. Information Discovery and Delivery, 52(2), 227–242. https://doi.org/10.1108/IDD-01-2023-0001

THE. (2023, June 12). Latin America University Rankings 2023. Times Higher Education (THE). https://www.timeshighereducation.com/world-university-rankings/2023/latin-america-university-rankings

Toledo, E. G. (2020). Why Books are Important in the Scholarly Communication System in Social Sciences and Humanities. Scholarly Assessment Reports, 2(1), 6. https://doi.org/10.29024/sar.14

Ullah, F., Shen, L., & Shah, S. H. H. (2023). Value co-creation in business-to-business context: A bibliometric analysis using HistCite and VOS viewer. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.1027775

Usher, A., & Savino, M. (2009). A Global Survey of University Ranking and League Tables. In The Routledge International Handbook of Higher Education. Routledge.

Vanz, S. A. de S., Dominique, A. P., Lascurain Sánchez, M. L., & Sanz Casado, E. (2018). Rankings universitários internacionais e o desafio para as universidades brasileiras. Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência Da Informação. Florianópolis, SC. Florianópolis. Vol. 23, n. 53 (Set./Dez. 2018), p. 39-51.

Velho, L. M. S. (2001). Estratégias para um sistema de indicadores de C&T no Brasil. 13.

Vieira, E. S., & Gomes, J. A. N. F. (2009). A comparison of Scopus and Web of Science for a typical university. Scientometrics, 81(2), 587–600. https://doi.org/10.1007/s11192-009-2178-0

Vieira, G. A., & Leta, J. (2024). biblioverlap: An R package for document matching across bibliographic datasets. Scientometrics, 129(7), 4513–4527. https://doi.org/10.1007/s11192-024-05065-5

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

Waltman, L., & van Eck, N. J. (2013). A systematic empirical comparison of different approaches for normalizing citation impact indicators. Journal of Informetrics, 7(4), 833–849. https://doi.org/10.1016/j.joi.2013.08.002

Zhang, C.-T. (2009). The e-Index, Complementing the h-Index for Excess Citations. PLOS ONE, 4(5), e5429. https://doi.org/10.1371/journal.pone.0005429

Zhu, J., & Liu, W. (2020). A tale of two databases: The use of Web of Science and Scopus in academic papers. Scientometrics, 123(1), 321–335. https://doi.org/10.1007/s11192-020-03387-8