Lost in citations: Why standard metrics fail archaeology and regional scholarship. A critical analysis of bibliographic indexing and research evaluation.

Andrea Di Renzoni

Institute of Cultural Heritage – National Research Council (Rome- Italy)

DOI: 0.32028/exnovo-vol-9-pp. 55-98

Abstract
This article explores the development, structure, and implications of bibliographic indexes and research metrics in contemporary academic publishing. It traces the historical evolution of indexing services from early systems like Index Medicus to modern platforms such as Web of Science, Scopus, Google Scholar, Dimensions, and OpenAlex. The study highlights how varying inclusion criteria, ranking algorithms, and citation policies among databases produce significantly different representations of scholarly output. It critically examines key metrics (e.g., Impact Factor, h-index, CiteScore, SNIP, SJR), assessing their impact on journals, authors, and institutions. It shows how current evaluation systems, including issues of classification and disciplinary bias, does not reflect the quality and impact of research outputs in the Humanities and Social Sciences, and especially in Archaeology. This is particularly evident in Italian prehistoric archaeology, explored here as a case study. The paper also addresses the ethical dilemmas linked to peer review, predatory publishing, and metric manipulation, advocating for more transparent, pluralistic, and context-sensitive approaches to scientific assessment. The final section reflects on the peculiarities of archaeological data and publication practices, underlining the inadequacy of conventional metrics for evaluating scientific contribution in this field.

Keywords: Bibliographic Indexes; Research Metrics; Citation Analysis; Scientific Publishing Ethics; Archaeology and Publishing

CONTACT: Andrea Di Renzoni , andrea.direnzoni@cnr.it

Introduction

Academic journals first appeared in 1665 with the publication of *Journal des Sçavans* and *Philosophical Transactions*. The former was edited by Denis De Sallo, who later founded the Académie Royale des Sciences; the latter by Henry Oldenburg, the first secretary of the Royal Society of London (Fyfe et al. 2022). Both societies quickly became central institutions in the development of modern science, fostering systematic inquiry and the public dissemination of knowledge.

Scientific papers quickly became the primary medium for sharing research, leading to a rapid expansion of academic journals. By 2007, an approximation based on Ulrich's database[1] suggested the existence of 23,750 peer-reviewed journals (Björk et al. 2008). This exponential growth is believed to have been largely driven by scientific specialization (Tenopir & King 2014: 161) that in turn led to the emergence of new journals dedicated to specific research fields and the subdivision of existing ones.

This growth was also accompanied by an increasing number of private publishers. While many early journals originated from scientific societies, commercial publishers began playing a key role as early as the Victorian era, proving more efficient in distribution. Over time, scientific publishing became an attractive market for commercial enterprises. By 2013, nearly half of all journals indexed in Web of Science were published by just five major publishing houses (Larivière et al. 2015: 5).

The expansion of academic publishing, both in the number of journals and the rise of commercial publishers, often unaffiliated with universities or research institutions, has necessitated efficient search tools and mechanisms to assess journal credibility. These needs have driven the development of journal indexing and ranking systems, which today have a significant impact on researchers' careers, institutional performance, and funding opportunities.

The discussion that follows addresses the main features of indexing systems, including their approaches to content selection, quality assurance, and research evaluation, while also considering the broader implications and criticisms associated with their use.

In the second section, I address the issue of indexing by introducing the most important databases in use. The third section analyzes the problems related to measuring scientific production, using Italian archaeology as a case study. This leads to the following fourth section, which discusses how the quality of research outputs is challenged by the push toward hyper productivity in the academic system. The final section considers the scientific process in archaeology and the need to ensure that different kinds of scientific products receive the recognition they deserve.

Indexing

Indexing is crucial for scientific journals, which constitute a fundamental part of scholarly communication. Without comprehensive and well-structured indexes, a vast portion of

---

[1] Ulrich's Periodicals Directory (ISSN 0000-0175, and ISSN 0000-2100) is the standard library directory and database providing information about popular and academic magazines, scientific journals, newspapers and other serial publications (definition by Wikipedia.org).

the academic literature would remain inaccessible to researchers. However, indexing periodicals presents some challenges:

- It is a continuous process, often requiring a team of indexers;
- Journal articles cover a broad and evolving range of topics;
- The terminology used for indexing must remain consistent over time and across disciplines;
- Multiple thesauri exist and selecting the most appropriate one can be challenging.

The birth of indexing

The need to organize knowledge has been recognized since antiquity. However, indexes in the modern sense, providing precise locations of names and subjects within a text, were not compiled in ancient times and remained rare before the advent of printing (Wellish 1983: 149).

A major advancement in periodical indexing came with William Frederick Poole, who, as a student at Yale and librarian of a college society, developed *Poole's Index to Periodical Literature*. This subject index (1802–1906) covered articles from 470 English and American magazines (Carlson 1928: 30).

In 1879, the *Index Medicus* was introduced by the Library of the Surgeon General's Office of the United States Army (now the National Library of Medicine). Conceived as "a monthly classified record of the current medical literature of the world", it was later complemented by the *Index-Catalogue*, published in 1880 following an initiative by John Shaw Billings. The systematic indexing of medical literature was considered by William Henry Welch, a leading pathologist and bibliophile, to be "America's greatest contribution to medical knowledge" (Greenberg & Gallagher 2009: 108). In 1971, *Index Medicus* evolved into the MEDLINE database, which was later integrated into *PubMed* (1997), now the most important online indexing system for health sciences.

The transformation of *Index Medicus* into a digital format illustrates how the "digital revolution" has profoundly reshaped scientific publishing. The rapid expansion of the internet has further accelerated the adoption of digital formats, altering how researchers search for and access information (King et al. 2009). However, the digital era has also amplified existing challenges in the indexing process, including maintaining consistency across taxonomies, adapting to evolving terminologies, and ensuring long-term accessibility of indexed materials.

The number of journals and articles & indexes

Many efforts have been made to determine the total number of journals and articles published over time. Jinha (2010) estimated that by the end of 2009, the total number of peer-reviewed articles had reached 50 million. As previously mentioned, specialization is often considered the main driving force behind the increasing number of journals. Some authors, such as Mabe (2003), identify a clear correlation between the number of researchers and the number of articles produced, which in turn influences the growth in journal titles. Conversely, Hanson et al. (2023: 3) argue that article growth results from a more complex interplay of factors, which they refer to as "the love triangle of scientific publishing".

This triangle consists of three key players: publishers, researchers, and funders (institutions). Publishers, as one vertex, seek to publish as many articles as possible while maintaining a certain quality standard. Researchers, another vertex, are primarily driven by the opportunity to publish in high-ranking journals, as publication and citation metrics are crucial for employment, promotion, and funding opportunities. Institutions, the third vertex, evaluate researchers competing for funds based on these quality metrics. This push-and-pull dynamic creates a self-reinforcing system in which all actors share a common concern: quality. However, defining quality remains a challenge. The previously mentioned *Index Medicus* provides a key example of the role of indexes in quality control. It has been argued that, compared to the *Index-Catalogue*, *Index Medicus* was significantly more selective in scope, concentrating on newly published articles from selected journals, as well as specific books and theses (Greenberg et al. 2009: 109). The term "selected" is particularly revealing. Today, journal ranking metrics are largely based on citations received by articles published in "selected" high-ranking journals included in specific indexing databases. With the growing importance of metrics, journal prestige is no longer based solely on the reputation of its editorial board or affiliated institutions. Instead, indexing status has become a central, and often dominant, determinant of perceived quality.

Indexing database

The history of indexes is extensive, and their importance to the research process is unquestionable. They help authors identify relevant sources, and indexed articles are more likely to be cited. In addition, indexes facilitate citation counting, which underpins (almost) all research metrics. A recent overview of existing databases has been published in *Archeologia e Calcolatori* (Di Renzoni 2025), while a more detailed list of databases and related services operating within the academic publishing ecosystem is provided here in the Appendix. This section offers only brief notes on three major indexes, selected for their relevance to the discussion that follows.

<u>Web of Science (WoS)</u>

In 1955, Eugene Garfield conceived an idea that led his Institute for Scientific Information (ISI) to create the Science Citation Index (SCI) in 1964. This system, based on organizing information through citation connections, anticipated web hyperlinking and the Google Search algorithm by three decades. ISI gradually expanded its indexing scope, adding Social Sciences (SSCI) in 1973 and Arts and Humanities (AHCI) in 1978. A significant development came in 1976 with the introduction of the Journal Citation Reports (JCR), which analysed journal-to-journal citations to evaluate the influence and prestige of specific titles and map the scientific communication network. Among the various metrics introduced, the Journal Impact Factor (see Appendix) became the most influential. The organization underwent several structural changes until it concretized in Clarivate Web of Science (WoS), and in the new reboot of ISI, in 2018 (Clarivate History, 10 Feb 2025).

Clarivate's description of WoS emphasizes three key characteristics: trustworthiness, publisher independence, and comprehensiveness. Their website states:

"*The Web of Science is the world's most trusted publisher-independent global citation database. […] [The] independent and thorough editorial process ensures journal quality, […] creating the most comprehensive and complete citation network to power both confident discovery and trusted assessment*".
This emphasis reflects the crucial role of citation metrics in contemporary academic publishing and highlights the importance of selecting journals based on reliable criteria independent of publishers.
The selection process for WoS journals follows rigorous standards. And, to ensure impartiality, the criteria designed to select journals are evaluated by in-house editors who maintain no affiliations with publishing houses or research institutes, thus avoiding potential conflicts of interest. (Clarivate Editorial Selection Process, 10 Feb 2025)

Elsevier Scopus
The emphasis that Clarivate places on the separation between publishers and indexers stands in marked contrast to Scopus, a widely used index developed and managed by Elsevier, one of the world's largest academic publishers. While Web of Science builds much of its credibility on being "trusted" and "publisher-independent", Scopus cannot make the same claim, as it is operated by a commercial publishing house. Instead, Scopus highlights its commitment to "transparency" and relies on an "independent board" of subject experts who continuously review new titles using both quantitative and qualitative criteria (Scopus Selection Criteria, 10 Feb 2025).
Since 2016, journals have been ranked according to the CiteScore metrics, a family of eight indicators designed to evaluate the influence of serial titles. Elsevier emphasizes the transparency and reliability of its system, stating: "Scopus metrics are a comprehensive, trustworthy and transparent way to demonstrate your journal, article, author and institutional influence" (Scopus Metrics, 10 Feb 2025).

Google Scholar (GS)
Google Scholar (GS) is a search engine for scholarly literature that indexes publications across multiple disciplines and sources. It operates similarly to Google Search, using automated software known as "crawlers" to discover and index academic documents. However, GS does not apply expert supervision or any formal quality control in selecting scientific content. It collects research papers from across the web, including grey literature and non-peer-reviewed articles and reports.
GS employs various ranking criteria for search results, papers, journals, and authors. While the exact ranking algorithm is not publicly available, some studies (Beel, Gipp 2009: 236) have attempted to analyse its mechanisms. The main factors influencing search result rankings include: 1) the number of citations an article has received; 2) the presence of search terms in the title (GS does not consider synonyms); 3) a relatively low weighting given to the frequency of search terms in the full text; 4) a preference for more recent

articles over older ones; 5) a strong weighting assigned to author and journal names, reinforcing the so-called "Matthew Effect"[2] .

For journal ranking, GS uses "Scholar Metrics", updated annually. The latest release (July 2024) covers articles published between 2019 and 2023 and includes citations from all articles indexed in GS as of July 2024. Journals are ranked based on the following metrics, calculated over the past five years: h-index (see Appendix); h-core, the set of articles that contribute to the h-index; h-median, the median citation count within the h-core; h5-index, h5-core, and h5-median, the same metrics calculated based on articles published in the last five complete calendar years.

Authors are evaluated using similar metrics, including the h-index and the h10-index.

Individual papers are ranked based on their citation count (e.g., GS's "Classic Papers" section highlights highly cited works). GS retrieves bibliographic data and citation links between papers using automated software called "parsers". Since references are identified algorithmically, without human verification or correction, GS does not guarantee complete accuracy. Errors in citation identification can result in missing papers, lower rankings in search results, or inaccurate metrics. Citation count discrepancies are not uncommon. To illustrate these shortcomings, I provide examples from my own GS profile: Paper A: 2 incorrect citations out of 19 total; Paper B: 1 incorrect citation and 1 duplicated citation out of 13 total; Paper C: 1 incorrect citation out of 19 total. These inaccuracies highlight the limitations of GS's automated citation indexing and ranking system.

Measuring Science: Insights on Metrics from an Italian Case Study

In the past decade, metrics' importance has become central to securing funding and, more generally, advancing academic careers. The focus on metrics is exemplified by the prominence given to journal scores displayed on their websites or by the fact that the excessive attention paid to these metrics has been labeled the "Impact Factor obsession" (Hicks et al. 2015). Numerous metrics exist to evaluate journals, authors, and individual scientific products such as papers, patents, and datasets.

Even if non-exhaustive, the review of the metrics discussed in the appendix highlights that almost every index uses citations as a proxy, which opens a range of potential pitfalls:

- Citation count depends on the databases used, each of which covers scientific output in different ways.
- Metrics (very often) do not distinguish between positive and negative citations; some scientific works may be cited due to controversy or errors.
- Practices for producing outputs vary significantly across disciplines:
- Some formats are less likely to be included in bibliographic databases, such as in the Social Sciences and Humanities, where books and locally relevant journals are crucial but often not indexed.

[2] The "Matthew effect", also known as the principle of accumulated advantage, describes the phenomenon whereby those who start with an advantage tend to gain even more benefits over time, while those who begin at a disadvantage tend to fall further behind.

- In some areas, researchers are encouraged to publish frequently, while in fields where monographs are important, such as in archaeology, publishing one book every few years might be seen as appropriate.
- While English is widely used, in some contexts, the national language may be preferred, which can reduce the likelihood that the work will be included in international databases.
- The number of authors per publication also varies by discipline, affecting the allocation of credit and the interpretation of bibliometric indicators.
- The way indexes are conceived can be misleading. For instance, the Journal Impact Factor (JIF), the most commonly used journal metric, has been criticized for its inability to properly represent individual articles (Seglen 1997).

At a more general level, some authors have raised profound concerns about the "bibliometric approach" to evaluating science. These concerns are eloquently captured in the title "Not everything that can be counted counts, and not everything that counts can be counted" (Olive et al. 2023). Authors argue that the evolution of journals has been significantly influenced by two interconnected forces: the information revolution and neoliberalism, leading to an increase in competition and managerial practices. Bibliometrics, initially designed to assist librarians in selecting journals for their collections, have now become central to the neoliberal university culture, heavily influencing academic evaluation, peer review, and promotion processes. As a result, metrics have increasingly driven the push for "hyperproductivity", which, while boosting publication numbers, may undermine the quality of research.

The thesis expressed by Olive et al. is part of a broader debate concerning how metrics should be used more effectively, and when they should be resisted altogether (Phillips 2020). These concerns have gained traction in various declarations, manifestos, and reports that argue for a critical approach to the use of metrics. Notable examples include the "Declaration on Research Assessment" (Dora, 10 March 2025) the "Leiden Manifesto for Research Metrics" (Hicks et al. 2015), and "Metrics Tide" (UKRI, 10 Mar 2025), all of which emphasize the importance of contextualizing metrics and caution against relying on them as absolute measures of scientific value.

The following simple tests aim to highlight the problem of using metrics without proper contextualization within specific research fields. Italian archaeology was chosen as a case study because it exemplifies a system positioned between globally recognized indexes and locally developed practices, and it is also the field I know best, which ensures appropriate handling of the data.

Index comparison

Taken together, the indexes discussed above suggest that the way they are structured, particularly the varying policies on publication selection and citation management, leads to significantly different outcomes (Martín-Martín 2021). The most substantial divergence lies in the inclusion criteria adopted by each database, which range from strict, quality-based selection to broader indexing of web-crawled content. The picture becomes more complex in disciplines where indexes hold some importance, yet bibliometric criteria are not officially adopted. This is the case for the Social Sciences and Humanities in Italy, where journal relevance is formally assessed by ANVUR, a governmental agency; nevertheless, the visibility and ranking of journals in international indexes still appear to inform evaluation practices, at least implicitly[3].

ANVUR

The National Agency for the Evaluation of the University System and Research (ANVUR) is the Italian authority responsible for assessing the quality of higher education and research in Italy since 2011. Among its quality control activities, ANVUR classifies scientific journals for the purpose of evaluating scholars' National Scientific Qualification (ASN[4]) indicators. However, ANVUR does not rank journals; rather, it classifies them as either "scientific" or as "scientific journals with Class A status" (ANVUR, 30 Jun 2025). A "Classe A" journal is a scholarly publication that meets the highest standards of scientific quality within a specific disciplinary sector and is recognized within the Italian system as a benchmark for academic research and evaluation in that field. Inclusion in the "Classe A" list has significant implications for individual career advancement and institutional assessments in the Italian academic system. To obtain the National Scientific Qualification (ASN), a scholar must publish a minimum number of papers in Class A journals. Accordingly, achieving Class A status is crucial for a journal to attract submissions and gain recognition.

To be included in ANVUR's journal "Classe A" list, a publication must undergo an evaluation process involving multiple actors. A prominent role is played by the Working Group (WG), which is appointed by the Governing Board (Consiglio Direttivo, CD) and composed of qualified scholars in the relevant scientific area, chosen from a list of eligible experts. The WG assesses whether a journal meets the criteria for inclusion in the Classe A category.

Within the Italian system, archaeology is part of "Area 10," which includes antiquities, philology, literary studies, and art history. Each area is further subdivided into Scientific Disciplinary Sectors (Settori Scientifico-Disciplinari, SSDs). The list of eligible experts, as of the 2024 call, included Full and Associate Professors from Italian universities (Fig. 1a), along with scholars from three foreign institutions, representing a wide range of disciplines. Figure 1b shows the distribution of eligible experts across SSDs, revealing the underrepresentation of L-ANT sectors, with only 5 out of 88 members.[5] This underlines the need for external experts, who are proposed by the WG and approved by the CD.

---

[3] Although difficult to measure, the prestige of certain internationally recognized journals clearly shapes readers' perceptions of individual papers, often beyond the intrinsic quality of the research, even when national systems like the Italian ANVUR assign formal journal classifications. This leads scholars to cite papers from these journals more frequently, a phenomenon that can be termed the "journal prestige effect".

[4] The National Scientific Qualification (*Abilitazione Scientifica Nazionale*, ASN) is a centralized evaluation process in Italy that certifies a scholar's eligibility to apply for associate or full professorships at Italian universities, based on the assessment of their scientific qualifications and research output.

[5] Until 31/12/2027, member of the WG of area 10 pertain to the following SSD: L-ANT/09, L-ART/04, L-ANT/05, L-FIL-LET/10, L-LIN/1

In addition, an important role in the journal evaluation process is played by assistants to the WG disciplinary area groups, chosen among the eligible expert. They are responsible for evaluating general requirements, facilitating communication between ANVUR offices and experts, and assisting external reviewers in carrying out their tasks.

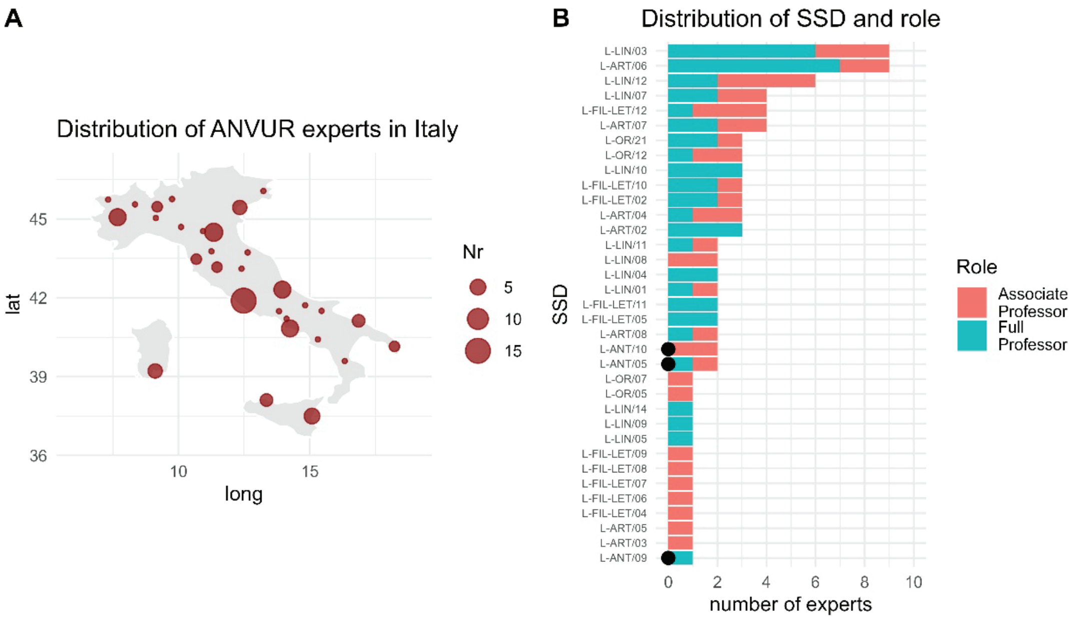

Figure 1. Geographic distribution of ANVUR eligible experts across Italy based on their affiliated universities; B: Distribution of ANVUR eligible experts by Scientific Disciplinary Sectors (Settori Scientifico-Disciplinari, SSD). Within Area 10, archaeology is represented by the following SSDs: L-ANT/01 (Prehistory and Early History), L-ANT/02 (Greek History), L-ANT/03 (Roman History), L-ANT/04 (Numismatics), L-ANT/05 (Papyrology), L-ANT/06 (Etruscology and Italic Antiquities), L-ANT/07 (Classical Archaeology), L-ANT/08 (Christian and Medieval Archaeology), L-ANT/09 (Ancient Topography), and L-ANT/10 (Methods of Archaeological Research), among others.

<u>Comparing metrics</u>

To better understand how the structural differences among indexes affect the representation of a specific discipline within a specific system, a comparative analysis was conducted focusing on the evaluation criteria for archaeological publications in the Italian system. I will demonstrate that the mechanical attribution of a value to a scientific article is, in many cases, misleading. The differences among the more commonly adopted databases were tested in two ways. First, the number of journals under the category "Archaeology" was retrieved from Scopus, WoS, Dimensions, OpenAlex, and Google Scholar. Data from Scopus and WoS were accessed via an institutional subscription, OpenAlex data were retrieved using its free API with a Python script, Dimensions data were queried through its web interface and copied into a spreadsheet, and Google Scholar results were manually extracted. Some clarifications are needed regarding how categories were defined in different datasets:

- *WoS* (as of 5 February 2025) lists 21,973 journals, classified into 254 categories, including "Archaeology", which contains 165 journals.
- *Scopus* (as of February 2025) lists 43,703 journals. Users can filter using the Subject Area field. "Archaeology" appears as a subgroup in both Social Sciences and Arts and Humanities (with minor differences: 463 journals under Arts and

Humanities, 409 journals under Social Sciences, and 496 journals when searching across both categories). The latter approach was used in this study.

- *Dimensions* allows users to filter papers by type and category. "Archaeology" is part of the broader field History, Heritage, and Archaeology. The platform aggregates journals based on the number of articles classified as archaeological, resulting in 4,501 journals after filtering.
- *OpenAlex* (as of February 2025) indexes 210,023 journals, each classified into multiple topics. A single journal may belong to multiple topics, with the number of associated papers available for each. Topics are structured into subfields, which belong to fields, which in turn belong to domains. Two subfields named "Archeology" (note the missing "a") exist, one under Social Sciences and another under Arts and Humanities. Data were retrieved via the OpenAlex API, and each journal was assigned to the subfield with the highest paper count, resulting in 2,430 journals classified as "Archaeological Journals", ranked by OpenAlex's h-index.
- *Google Scholar* does not provide an official API, but commercial APIs (e.g., SerpAPI) exist for search engine queries. Users can manually extract the top 20 journals in the Archaeology category, ranked by their h-index.
- *ANVUR*, the Italian National Agency for the Evaluation of Universities and Research Institutes, does not rank journals but categorizes some of them into "Classe A" journals, based on their quality. The latest list (late 2024) includes 2,314 journals in this category.

A comparison of the top 20 archaeological journals across these indexes reveals limited overlap. In total, 63 unique journals were identified across the five rankings. To compare them, a ranking score was assigned using the reversed ranked score sum, weighted by coverage (i.e., the proportion of databases that indexed each journal). The formula used is:

$$Ranking\ Score = \sum \left((Rank_{max} + Rank_{min}) - Rank_{journal} \right) \frac{nr\ indexes\ including\ journal}{5}$$

Notably: only 11 out of 63 journals appear in at least three indexes, and only one journal is present in all five indexes (Appendix, Table 1).

The similarity between indexes was analyzed using a modified Spearman correlation, described as follows:

- Compute the Spearman correlation (SC) between rankings.
- Normalize the correlation: $SC_n = (SC + 1)/2$ to constrain values between 0 and 1.
- Apply a weight: $SC_w = SC_n \frac{Nr\ of\ shared\ journals}{Max\ nr\ of\ shared\ journals}$

The resulting heatmap (Fig. 2) indicates that WoS, Scopus, and Google Scholar form a cluster of similar indexes.

One major issue affecting comparisons is the classification of journals. For example, "Forensic Science International" is categorized under Archaeology in OpenAlex due to the relevance of some articles, but Scopus and WoS classify it differently: Scopus: Social Sciences, Law, Medicine (Pathology and Forensic Medicine) WoS: Medicine (Legal). Such discrepancies can significantly affect rankings and impact metric consistency.

The classification issue also affects the comparison between the five indexes and the ANVUR list. Among the 63 journals identified in this study:

- 25 journals (≈40%) are not listed in the ANVUR Classe A category, some of which are clearly archaeological journals.
- Overlaps with ANVUR's list:
  - OpenAlex and Dimensions share 15 out of 20 journals with ANVUR.
  - Google Scholar shares 14 out of 20.
  - WoS shares 13 out of 20.
  - Scopus shares 12 out of 20.

Interestingly, the two most selective indices (WoS and Scopus) exhibit the lowest degree of overlap with ANVUR, despite the fact that all three are formally grounded in quality-based criteria. This highlights how different evaluation methods can produce divergent outcomes. Many journals that focus on local topics or specific chronological frameworks are rarely indexed in the major international databases, whereas a system like ANVUR's takes these specificities into account by relying on experts familiar with the national academic and disciplinary context.

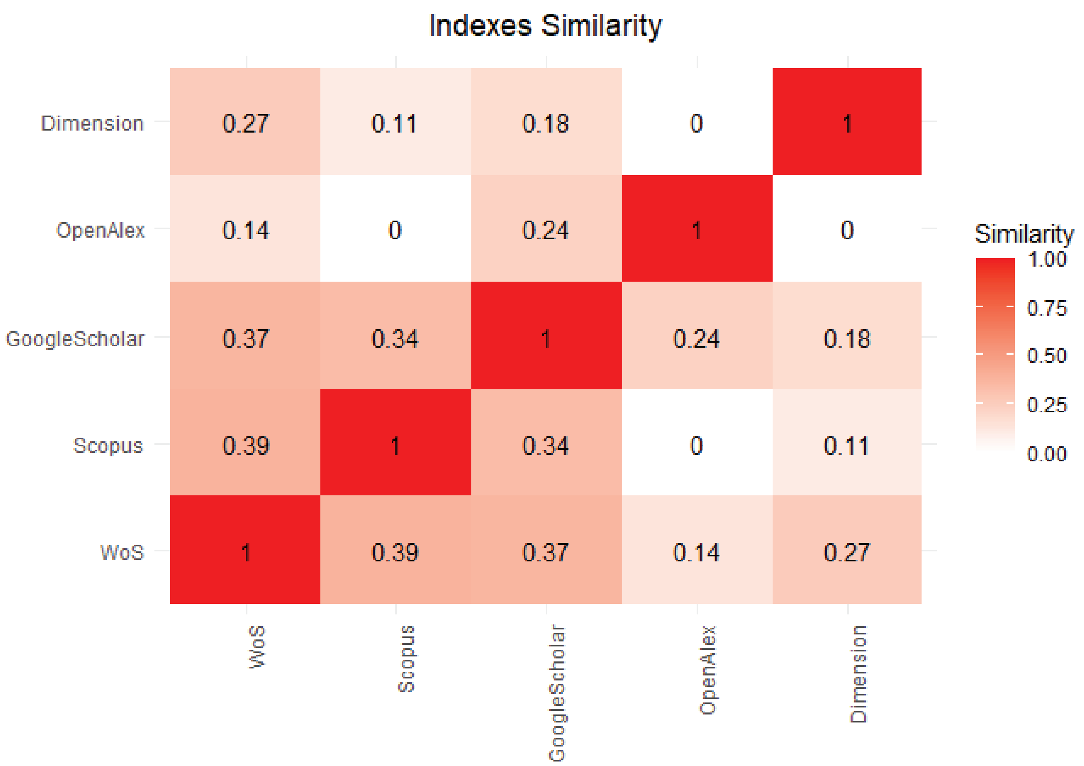

Figure 2. Heatmap of similarity across bibliographic indexes.

Testing Citation databases

The second test consists of a simple experiment comparing the citation count of the same paper across different databases. As first example I will use a paper I co-authored in 2019 published in *Scientific Reports*, a well-known journal indexed by many databases. Scopus reports 40 citations, WoS 31, Google Scholar 56, OpenAlex 46, Dimensions 40, Lens 38, CrossRef 36, and ResearchGate 49.

As a second example, I selected a paper regarded as seminal for the Middle Bronze Age in Central-northern Italy (the Terramare culture) published in 2009 in Italian in "Scienze della Antichità", a journal published by Sapienza University of Rome, that therefore is rarely indexed by the main databases, but categorised as Classe A in the Italian system). The results were striking: the article is not indexed in WoS, Scopus, OpenAlex, Lens, Dimensions, or CrossRef. However, Google Scholar reports 163 citations, and ResearchGate reports 104. This simple experiment shows two great limitations of citations count: first, the importance of a paper is not directly linked to the journal in which it is published; a journal could not be indexed in the main databases but, the same, could be of great importance for some branches of the discipline it pertains to.

The issues with metrics in the humanities are well documented by Emanuel Kulczycki et al. (2018). Authors point out that traditions, publication patterns, and language strongly influence the reliability of citation indexes. In Italy and beyond, in the Arts, Humanities, and Social Sciences (AHSS), researchers produce a wider variety of outputs, such as books and book chapters, compared to the primarily journal-based outputs in other disciplines. Many of these output types are not well represented in databases, or their citation records are incomplete, making meaningful analysis difficult. Additionally, language barriers and the specific focus of research topics in AHSS fields contribute to the inconsistency of citation indexes.

They found that in many European countries, less than 50% of Social Sciences and Humanities publications are visible in WoS. Authors in these fields often choose to publish in journals with a narrower geographical and thematic scope but relevant for their discipline and country, that cater to specific sectors within AHSS fields, which can result in their work being underrepresented in global citation databases. This is particularly evident when comparing the ANVUR list with the top 20 archaeological journals identified by major indexes and the results of the simple test described above, where an important journal for Italian archaeology, such as *Scienze dell'Antichità*, is perceived as less significant than it truly deserves. Prestigious international journals often privilege broad, "catchy" themes over studies rooted in newly excavated or contextual data. Detailed datasets may be sidelined in favor of general interpretations, making it harder for context-rich research to gain visibility, while topics with wider appeal attract disproportionate international attention.

DoRA & the Leiden Manifesto

Concerns about the unquestioned use of metrics have long been felt across disciplines. The topic has been addressed by scientometricians, leading to the formulation of two important declarations. In 2012, at the Annual Meeting of the American Society for Cell Biology in San Francisco, the Declaration on Research Assessment (DORA) was developed. DORA recognized the urgent need "to improve the ways in which researchers and the outputs of scholarly research are evaluated." It emphasized the pressing need to reform how the output of scientific research is assessed by funding agencies, academic institutions, and other stakeholders. Among its key claims were: the need to eliminate the use of journal-based metrics, such as Journal Impact Factors (JIF), in funding, appointment, and promotion decisions; the need to evaluate research based on its own merits, rather than the journal in which it is published; and the need to take advantage of the opportunities offered by online publication

In 2015, the journal "Nature" published a comment paper titled "Bibliometrics: The Leiden Manifesto for Research Metrics" (Hicks et al. 2015), in which the authors advocated for the use of ten principles to guide research evaluation. The authors, who were scientometricians, social scientists, and research administrators, expressed growing concern about the widespread misuse of indicators in the evaluation of scientific performance. They observed how the abuse of research metrics had become too widespread to ignore and, in response, presented the Leiden Manifesto.

The emergence of these initiatives is highly significant. Over a decade ago, the "obsession with the JIF" was already identified as a problem affecting researchers' attitudes towards science and, by extension, the quality of science itself. This international debate is reflected also in Italy.

The quality dilemma

The phrase "publish or perish", as we understand it today, was first coined in the 1920s by Clarence Marsh Case (Moskovkin 2024). While every scholar today is familiar with its meaning, few are aware of its origins in a Jesuit proverb, which stated: "publish lest the knowledge should perish with you" (Seppelt et al. 2018). This original phrase carried a surprisingly different message from the modern interpretation, which is associated with the pressure researchers face to publish frequently, often at the expense of quality.

In a recent talk for the opening of the 2023 Academic Year at the "Accademia dei Lincei" in Rome[6], influential Italian archaeologist Marcella Frangipane called attention to two critical aspects of research (Frangipane 2023). First, she highlighted the meaning of "innovation," which is often equated with technical progress but should instead be understood as the "ability to look beyond mainstream issues and the current 'priorities' of the moment". Second, she emphasized the need for "time" in the research process. Frangipane aligns with the "Slow Science Movement", which opposes performance targets and advocates for research driven by curiosity rather than short-term productivity. She cites a recent article published in "Nature" (Park et al 2023), which analysed the "disruptiveness" of scientific papers and patents over time. Using the CD5 index, the authors found that while the number of disruptive articles remained relatively stable, the average breakthrough capacity had significantly declined. This trend suggests an increase in "background noise", irrelevant or marginally relevant articles. Moreover, scientists are

[6] The Accademia Nazionale dei Lincei, founded in 1603, is one of the oldest scientific academies in Europe. It promotes the advancement of knowledge across the sciences and humanities and brings together leading scholars at both national and international levels.

increasingly relying on a narrower range of existing knowledge, struggling to keep up with the pace of scientific production. The authors concluded that rather than a fixed “carrying capacity” for highly disruptive science, the shift towards quantity over quality confines researchers to familiar, smaller areas of knowledge. This benefits individual careers but does little for scientific progress as a whole.

In his 2017 article, “Is the staggeringly profitable business of scientific publishing bad for science?” (Buranyi 2017), Stephen Buranyi explores how commercial interests have shaped the field of scientific publishing. He draws attention to the industry’s extreme profitability, with margins surpassing even those of tech giants, due to the “triple-pay system” as described in a Deutsche Bank report. Governments fund research, scientists, whether as authors or reviewers, and institutions repurchase the final products at high costs. This business model, which traces its origins to Robert Maxwell’s Pergamon Press, revolutionized scientific publishing in the 1950s.

These observations are echoed by Young et al. (2008), who identify several consequences of the publishing ecosystem described above: 1) Articles published in highly competitive journals tend to present exaggerated results; 2) a small number of journals determine the visibility of most scientific discoveries; 3) scientific “herding,” or following the leader, compels authors to pursue popular research topics, often neglecting innovative ideas and independent paths of inquiry; 4) the artificial scarcity created by extremely low acceptance rates signals status, even if the content of the articles is not truly groundbreaking; 5) branding, wherein publishing in selective journals serves as evidence of a research result’s value, independent of its actual merit, becomes a key factor in career advancement.

In this environment, the long, slow, and nearly directionless work pursued by influential scientists like Fred Sanger, who published very little between his Nobel Prizes in 1958 and 1980, has become virtually unviable. Today’s system would likely have left Sanger without a position, despite his groundbreaking contributions to science (Buranyi 2017).

These structural traits give rise to serious ethical concerns, both for researchers and publishers. The concept of “quality” in research has become increasingly elusive, and its control is entrusted solely to the self-correcting mechanisms of the peer review process, which, however, is not without its flaws.

Peer reviewing

The peer review process (PRP) has long been considered the “self-correction mechanism” of science, ensuring the quality of published research since the advent of scientific journals. Peer-reviewed journal articles are seen as reliable because they have undergone independent evaluation by experts in the field. However, the confidence placed in this system appears to rest on shaky foundations, as several critical aspects of the process have been raised. In 1985, Stephen Lock, then editor of the “British Medical Journal”, published an entire book analysing the PRP, titled “A Difficult Balance: Editorial Peer Review in Medicine” (Lock 1985a). Subsequent events, including thematic congresses held by the Journal of the American Medical Association in 1989, 1993, and 1997, as well as numerous books (e.g., Godlee, Jefferson 2000) and a plethora of scientific papers, have continued to scrutinize the PRP. One example of ongoing attention to the subject is the “PEERE” Cost Action (European Cooperation in Science and Technology), which ran

from 2014 to 2018 with 31 participating countries, aimed to improve the efficiency, transparency, and accountability of peer review (PEERE, 14 Mar 2025).
Criticisms of the PRP stem from various factors:

- The definition of the PRP remains ambiguous (Smith 2006), as its mechanisms and objectives, whether for paper selection or quality improvement, are often unclear.
- The process is highly subjective, leading to a lack of consistency and transparency.
- It is susceptible to bias, including parochialism and misconduct.
- The process is costly in terms of the researchers' time, which can be quantified in economic terms (Aczel et al. 2021).
- Data regarding the PRP are generally inaccessible (Squazzoni et al. 2017).

As Smith succinctly puts it, "the practice of peer review is based on faith in its effects, rather than on facts.".
The "quality control" role of the PRP has often been questioned. Ozonoff (2024: 3) argued, "If peer review were a research instrument, we would be very reluctant to use it" and pointed out that the evidence linking pre-publication peer review to improved quality is, at best, mixed. Studies have also shown that the reviewers'ratings do not correlate with subsequent citations of the paper (Rangone et al. 2012, Bartneck 2017). Many experiments have been conducted to test this aspect of the PRP. For example, Smith (2006) describes an experiment conducted by the editorial board of the "British Medical Journal", in which major errors were deliberately inserted into a set of papers undergoing peer review. None of the reviewers detected all the errors, and most identified only a quarter of them. The list of "academic hoaxes" or bogus papers submitted to peer-reviewed journals is long. It ranges from the famous "Sokal Affair" (Sokal affaire, 15 Mar 2025) to the "Conceptual Penis" hoax (The conceptual penis, 15 Mar 2025), non-sensical articles that were published in peer-reviewed journals.
A significant concern regarding the PRP is its inconsistency and subjectivity. Reviewers agree only marginally more than random chance would predict. A famous experiment to assess the consistency of the PRP was conducted at the 2014 "Neural Information Processing Systems" (NIPS) conference in Montreal. To evaluate 10% of the conference papers (166 in total), the scientific committee was divided into two groups, each reviewing the same papers. The acceptance rate was set at 22.5%, but only 16 papers were accepted by both committees, meaning that more than half of the accepted papers were rejected by at least one of the committees. This high rate of disagreement (77.5%) suggests that the decision-making process may be closer to random than to a method based on expert evaluation. Among other factors, Brezis and Birukou (2016) attributed the arbitrariness to two main causes: 1) reviewers' preferences for similar ideas (homophily), and 2) differences in the amount of time reviewers allocate to evaluations.
Biases in peer review have been documented since the 1980s, including those based on author rank, gender, institutional affiliation, and research attitude (Peters, Ceci 1982; Lock, 1985a). Several studies have shown that even groundbreaking research, later awarded the

Nobel Prize, struggled to find publication due to the conservative nature of scientific communities (Campanario 2009).

Lock (Dean, Flower 1985: 1560) recalled an instance when he, as editor, responded to an author whose paper had been rejected by reviewers. He acknowledged that peer review "favours unadventurous nibblings at the margin of truth rather than quantum leaps" and suggested an experiment: to publish the paper along with all related correspondence and reviewers' reports, so that readers could better appreciate the editorial process. This idea of revealing the peer review process was put into practice by the "British Medical Journal" in 1999, where reviews were published online alongside the authors'original versions and responses to reviewers' comments (Smith 1999).

The issue of transparency in peer review has gained increasing attention. Open peer review, where review reports and reviewers' identities are published alongside the articles, has become a growing component of open science (Wolfram et al. 2020). Ozonoff (2024) argues that "real peer review happens after publication", as the scientific community continues to evaluate publications through citations, usage, contradiction, or disregard. This view is echoed by Oransky and Marcus (2011) that highlight the importance of post-publication analysis as part of the scientific record.

The rise of preprints and pre-reviewing platforms has also contributed to the evolution of the peer review process. Repositories like arXiv (arXiv, 20 Mar2025) allow authors to upload their works for public access and feedback before formal peer review. The Peer Community In (Peer Community, 18 Mar 2025), provides a platform for the evaluation and recommendation of research by the broader community. In the field of archaeology, the "Peer Community in Archaeology" (Peer Community Archaeology, 18 Mar 2025) offers a similar service, recommending noteworthy unpublished articles and enhancing their reliability through peer review, without the need for traditional journal publication[7]. The recommendations are published alongside all relevant editorial correspondence, including the reviews, the decisions of the recommenders, and the authors'responses. What is worth noticing is, more than the reviewers' names, making the review process transparent by showing what the reviewers noted about the paper, the authors' responses, and the editors' reasoning behind their decisions—thus providing readers with the full background of the final published outcome.

Ethics Issues

Academic Predators

The term "predatory journal" was first coined by librarian Jeffrey Beall in 2010 (Beall 2010, 2012), and since then, a significant body of literature has developed to address the issue and its potential solutions. The narrative about Predatory publishing is closely tied to the Open Access (OA) model, particularly the "author pay" system, neglecting the fact that the emergence of predatory journals is closely linked to the academic publishing

[7] The managing board of PCArc is composed of 11 members, including 1 Italian scholar. Among the 245 recommenders (those who manage preprint evaluations and play a role very similar to that of a journal editor), 19 scholars are affiliated with Italian institutions. The countries with the largest number of recommenders are France (32), Germany (30), the UK (23), Italy (19), and Spain and the USA (16 each).

system as a whole. OA aims to democratize access to knowledge and encourage wider dissemination and impact across the global scientific community. However, some authors believe that OA, particularly the Gold OA model, opens the door for dishonest publishers to exploit the "publish or perish mentality" for profit. Jeffrey Beall has been one of the most vocal critics of this practice. He authored the blog "Scholarly Open Access", which hosted his well-known "Beall's List", a blacklist of unethical journals he identified. His (controversial) views are outlined in his article "What I Learned from Predatory Publishers" (Beall 2017), written after he shut down his blog due to (alleged) pressure from his employer, the University of Colorado Denver, and fears for his job (the University of Colorado published a notice negating Beall's insinuations). Beall's criticisms of the OA movement are evident in his writings, where he describes the conflict of interest inherent in the author-pay model ("The more papers they accept and publish, the more money they make, meaning there is an ongoing temptation to accept unworthy manuscripts to generate needed revenue", 2017: 275) as a major threat to science ("I think predatory publishers pose the biggest threat to science since the Inquisition", Beall 2017: 276). He contrasts OA with the "old" subscription model, which he sees as more trustworthy, with reputable publishers ensuring the quality of articles. However, Beall's argument is flawed in several respects. First, there is no universally accepted definition of a "predatory journal" (Cobey et al. 2018). As Kyle Siler argues, economic exploitation can exist in various business models, and the term "predatory" is often subjective and context-dependent (Siler 2020).

In 2018, Amaral (2018) argued that comparing the behaviour of "predatory journals" to "traditional" publishers (he cites as an example Elsevier) is like comparing zooplankton to sharks, both in terms of scale and greed. While high fees in renowned journals may be criticized, they are rarely labelled as "predatory", even if their business models can be economically exploitative. This raises the question of where academic and professional gatekeepers should draw the line between legitimate and predatory publishing along this broad spectrum (Siler 2020). These findings underscore the ambiguity in academic publishing, where inequalities can exist both between and within publishing institutions, and where strong articles are published alongside more questionable content.

Beall's List, still accessible at beallslist.net, and similar initiatives aim to combat predatory publishing by compiling blacklists and whitelists. An example is Predatoryjournals.org, managed by anonymous volunteer researchers who have been affected by predatory publishers and seek to help others identify trustworthy journals. One of the most prominent, and controversial, blacklisting services is offered by Cabell Publishing, which provides a paid database of deceptive journals and a separate list of verified ones (Cabell, 18 Mar 2025). However, it has been criticized for inconsistencies in its evaluation procedures (Dony et al. 2020; Grudniewicz et al. 2019).

On the other hand, the Directory of Open Access Journals (DOAJ) is a widely respected website that maintains a community-curated list of reputable open access journals. Launched in 2003 and managed by Infrastructure Services for Open Access (IS4OA), DOAJ aims to increase the visibility and accessibility of high-quality, peer-reviewed open access journals across the globe, regardless of discipline or region.

Although some argue that the issue of predatory publishing is overstated (Olijhoek, Tennant 2018), the problem undeniably exists, though it should not be attributed solely to OA. As discussed in earlier sections, the structural issues in academic publishing are at the root of the problem. Numerous experiments (e.g., the "Dr. Fraud" experiment or the "Bannon experiment") have tested the vulnerability of the system, demonstrating that "suspect" journals are more prone to unethical behaviour than "controlled" journals, which, however, are not always able to guarantee quality control. A central issue in the debate is the role of peer review. Beall argues that in order to compete in a crowded market, legitimate OA publishers are pressured to promise shorter submission-to-publication times, which weakens the peer-review process. However, as discussed, peer review itself is often an opaque process, with the reputation of the journal serving as a proxy for quality and trustworthiness.

Ultimately, it seems that blacklists and whitelists are not foolproof tools but rather symbolic measures, tools we wish to believe in rather than ones that effectively address the underlying issues. In his 2018 article, Amaral provocatively claims that publishing in a high-impact factor journal is a collective illusion, promoted by funding agencies, institutions, and researchers, and that it serves as an excuse to delegate the evaluation of science to for-profit companies and anonymous reviewers, undermining objectivity. In this complex environment, researchers and institutions often rely on blacklists, whitelists, and committee guidelines without questioning the integrity of those gatekeepers, "Who watches the Watchmen?" (Strielkowski 2018). Academic publishing is both a professional and economic activity, and for it to be perceived as legitimate, a balance must be struck between these often-conflicting ideals (Siler 2020). The drive for profit by publishers is not new, as exemplified by the story of Robert Maxwell, and is not the only factor pushing science away from its meritocratic ideals. The history of scholarly publishing has long been shaped by power structures that favoured "whiteness, cis-gendered heterosexuality, wealth, the upper class, and Western ethnocentrism" shaping who was published and whose ideas were heard (Swauger 2017). These structural inequalities have historically marginalized voices from non-Western contexts and underrepresented groups, both in terms of authorship and editorial power. As a result, academic legitimacy has often been tied to institutions and journals rooted in the Global North, reinforcing epistemic hierarchies that still persist. In this context, predatory OA publishers are often associated with publishers from the global south (Beall 2012). However, the role of some journals in promoting research on regionally significant topics should not be overlooked (Cobey et al. 2018: 30).

<u>Thinner Slices, More Papers: "Salami Publishing", Plagiarism, and Self-Plagiarism</u>

Unethical practices are not limited to publishers alone; authors can also engage in questionable behaviour. Some of these practices have become so widespread that they have been the subject of specific studies and publications.

Plagiarism is one of the most common and perhaps the most pervasive form of misconduct on the part of authors. It can take various forms. As forensic plagiarism investigator Barbara Glatt describes (Lawrance 2024) it, plagiarism can be: Direct, copying word-for-word without attribution; Indirect, the wholesale theft of ideas; Mosaic, altering some words while copying others; Honest mistakes, unintentional errors of omission or execution.

Academic plagiarism often involves power dynamics, such as those between professors and students or across gender lines, where individuals in positions of authority may exploit their power to appropriate the work of others without proper credit. One of the more subtle forms of plagiarism is what I refer to as "secondary plagiarism", which operates at a structural level. This occurs when credit for a scientific idea, despite the work being published with all authors acknowledged, is attributed to a single author. "Secondary plagiarism" could have a much greater impact on researchers' careers (and mental well-being) than one might imagine.

Plagiarism is deeply ingrained in academic culture, with even prominent scholars being accused of it (e.g., former Harvard President Claudine Gay, Lawrance 2024). Several studies have sought to quantify the extent of plagiarism, with findings that suggest between 3% and 7% of scholars admit to having stolen ideas at least once. Even more striking, 30% admit to knowing colleagues who have plagiarized (Pupovac, Fanelli 2014; Xie et. al 2021; Allum 2024; Brooker 2024). This problem may intensify with the widespread use of generative AI tools, which have made rephrasing easier than ever. More concerningly, authors who misuse such tools may inadvertently plagiarize, often without understanding whom they are plagiarizing or realizing that they are doing so. ChatGPT, for example, is a widely used language model designed to create original outputs, but it does not actively verify whether the generated text matches existing sources, and, as a result, it may unintentionally reproduce commonly cited phrases or well-known passages. (OpenAI, 18 Mar 2025).

The increasing prevalence of plagiarism checkers underscores the widespread nature of this issue. For example, CrossRef offers the "Similarity Checker", which helps members prevent scholarly and professional plagiarism. Other services, such as "Grammarly" and "Scribbr", are also commonly used to detect and prevent plagiarism.

While plagiarism is generally recognized as unethical, what happens when an author plagiarizes their own work? Scholars like Vesna Šupak-Smolčić (e.g. 2013) have written extensively on self-plagiarism, arguing that it can artificially inflate an author's productivity. In response, the "Committee on Publication Ethics" (COPE) has issued guidelines for handling self-plagiarism. Šupak-Smolčić and Bilić-Zulle (2013), following Miguel Roig, classify self-plagiarism into four types, as shown above in table 2.

Research by Wager et al. (2015) highlights that highly prolific authors may publish more than one paper every 10 working days! In Archaeology and the Social Sciences, the most common forms of self-plagiarism are duplicate publications and text recycling, along with self-repetition. However, self-repeat plagiarism can be transparent and honest or deceptive, depending on the context.

Type of plagiarism	Description
Duplicate publication	When manuscripts are nearly identical, often involving the publication of the same article in different languages (some plagiarism checkers, such as "Turnitin", can detect these duplicated translations)
Salami slicing	A form of redundant publication common in experimental disciplines. This occurs when multiple papers are published from the same dataset or experiment, but each paper presents only a fraction of the overall findings. These papers may share similar hypotheses, methodologies, or results, but differ in text composition, making them harder to detect by software.
Augmented publication	Similar to salami slicing but involves adding new data to previously published work. These additions may appear modest but can artificially increase the number of publications based on a single study.
Text recycling	The simple reuse of previously published text, often through copy-pasting.

Table 2. Classification of plagiarism types, as defined by Šupak-Smolčić and Bilić-Zulle (2013) based on Miguel Roig's framework.

The factories of fake science: Paper Mills, Citation Mills, and the Industrialization of Academic Fraud.

Among the most unethical practices in scientific publishing, paper mills and citation mills are undeniably the worst offenders.

Paper mills are defined by COPE as "profit-oriented, unofficial, and potentially illegal organizations that produce and sell fraudulent manuscripts (containing fake and/or plagiarized data) that mimic genuine research. They may also handle the submission of articles to journals for review and sell authorship to researchers once the article is accepted for publication. Indications that manuscripts may have been produced by a paper mill are more apparent at scale, as they often share similar layouts, experimental approaches, and identical or altered images and figures (COPE paper mill, 18 Mar 2025).

A recent estimate suggests that at least 400,000 papers published between 2000 and 2022 may have been produced by paper mills, with only 55,000 of them being retracted, according to "Retraction Watch" data (Candal-Pedreira et al. 2022). Alarmingly, papers produced by paper mills are often widely cited, which complicates efforts to identify them. These papers can appear perfectly legitimate to reviewers and editors, but image analysis tools can sometimes detect image manipulation or duplication, though this remains a significant challenge. Software like "Problematic Paper Screener" may identify unusual phrases that hint at scientific misconduct. COPE has also published a list of common indicators to help identify suspicious papers.

Paper mills are most commonly associated with open-access journals, but they also infiltrate journals indexed in major databases Candal-Pedreira et al. (2022). However, the key issue lies on author side. As Cameron Neylon (2015) points out, researchers act out of rational choice within the competitive environment they themselves helped to create.

Review mills & Citation mills. In addition to paper mills, review mills and citation mills are emerging as related unethical practices. Review mills generate fake reviews using vague and repetitive formulas, undermining the peer-review process. Similarly, citation mills manipulate citation counts, often pressuring authors to cite specific articles. A study by Ibrahim et al. (2025) demonstrated how citation manipulation occurs. They found groups of scientists who received large volumes of citations from specific papers, raising concerns about the authenticity of these citations. In some cases, a paper consisting of just two pages would reference a particular author 29 times in its bibliography, with the main text containing only one citation. Authors even contacted “citation-boosting services” that sold batches of citations for prices ranging from $300 for 50 citations to $500 for 100 citations. Illicit citation sellers assured authors that the citations came from peer-reviewed journals indexed in Scopus, including some published by well-known houses like Springer and Elsevier, with impact factors as high as 4.79. Notably, the effectiveness of these citation-boosting services decreases dramatically when journals indexed in Scopus or WoS are involved.

A Virtuous Approach: Addressing Scientific Malpractices through Openness and Community Involvement

The scientific community is well aware of the dangers posed by malpractices, as evidenced by the growing body of research dedicated to the topic. While there are no simple solutions, and merely listing the good and bad practices does little to resolve the issue, several approaches show promise. One such path involves fostering openness in scientific products and encouraging wider participation from the scientific community in the “quality control” process. Several initiatives and models of scientific interaction are worth highlighting in this regard.

PubPeer is one such platform, created in 2012 to enable its community to discuss and review scientific research post-publication (PubPeer, 20 Mar 2025). Users can interact with one another and with the authors themselves, who have the opportunity to respond to comments. A study by José Luis Ortega (2022) analysed 17,244 commented publications. The study found that 12,687 (approximately 73.6%) of these publications showed signs of data manipulation or publishing fraud. Of these, 21.7% received editorial notices, and many were subsequently retracted. Notably, articles from Social Sciences and Humanities were less frequently flagged for issues (3.2%) compared to other disciplines.

An emerging model of scientific publishing is the **overlay journal**, which combines comments and published material in an innovative way. Overlay journals are a type of open-access academic journal that does not produce its own content but instead selects from texts that are already freely available online. Editors may formally republish the article with an explicit approval statement, add a note to the text or its metadata, or simply link to the article through the overlay journal’s table of contents. Another approach involves grouping scattered articles together into themed issues, which allows for a focused exploration of relatively obscure or newly emerging topics.

Indexes and (what matter in) Archaeology

In an always growing body of publications, indexes absolve three main functions: 1) they help researcher in gathering information about their own scientific field; 2) they act as guarantor of quality; 3) they measure scientific world actors (scholars, journals, institutions). It is out of doubt their immense utility and, at the same time, they are part of a system that should foster scientific advances but, instead, it is thought to be an obstacle to it because of the over production of poor significative papers, the homologation of ideas, the pushing through unethical and even illicit behaviours. It is beyond my competence and capacity to discuss how the scientific world should be organized but some consideration can be given to. The first function of the indexes, helping to search for data, does not need to be explained, we have just to keep in mind that the result of our searching depends on where we are searching, and that the selection operated by some indexes, that we can define the "trusted indexes" (namely Scopus and WoS that are considered in many institutional evaluation of scientists) can effectively be considered, to a certain extent, a quality control (as the firewall function against Citation Mills seem to indicate). Function 3 is a more complex point to discuss. As it has been stated in the preceding paragraph, metrics have a profound effect on scholar's careers in many disciplines. Metrics quantify quality largely based on citations. Thus, we should agree on what is relevant citations to be counted, what is meaningful for the advance of a discipline. It could be an endless task addressing this topic, we can therefore just consider some issues about Archaeology.

The nature of data and data-collecting in Archaeology

Data collection is the process of systematically gathering and measuring unorganized facts or figures. This data enables researchers to test hypotheses or develop new ones. Once analysed and possibly combined with other data, it transforms into meaningful "information", providing context and insight. Although methods vary across disciplines, accurate and honest data collection is universally essential.

In archaeology, however, the nature of data collection becomes more complex. What constitutes "raw data" or "interpreted data" can sometimes be difficult to distinguish, especially depending on the scale of observation. Modern excavations are typically organized around the concept of Stratigraphic Units (SUs) (or more in general layers or contexts), whose definitions are highly subjective and require specific expertise (De Guio 1988). The SU (in the form of maps and compiled sheets) enter in the archaeological literature as data but they can be considered "type 1 information", a product of an analytical process. Similarly, cultural objects found during excavation, such as pottery, also undergo a similar transformation. A sherd, the data we use in building chronological sequences, quantification, distribution, etc., can be considered the result of an analytical process rather than raw data. Its definition, which in most cases passes through the process of "archaeological drawing" and formalized description, that is how the "raw data" is presented in the literature, is the result of a subjective operation depending on the expertise of the individual archaeologist and can therefore considered "type 1 information".

In broader archaeological analysis, what is used are not samples of the "real world", but representations of it, produced through the analytical activities of researchers.

In archaeology, monographs are often the best, and sometimes the only, format for publishing archaeological data or "type 1 information". The production of these publications is time-consuming, as they result from extensive fieldwork and long-term data elaboration, and too often are considered merely descriptions, catalogues or lists of objects. This makes them poorly suited for the fast-paced "publish or perish mentality" prevalent in other disciplines.

Recent years have seen a shift towards more quantitative approaches in the Humanities and Social Sciences, often inspired by the "hard sciences" (defined by the use of mathematic). The rise of data modelling techniques, exemplified by the increasing use of programming languages like Python or R, marks a significant step forward for the field.

Consider mobility studies in archaeology. Researchers employ various analytical methods, from network analysis to complex statistical models, to reveal patterns of movement and interaction. But what exactly constitutes the foundational data for these analyses? In network analysis, archaeologists often map connections between sites (nodes) based on shared material culture (edges). Yet this seemingly straightforward approach conceals layers of interpretation. What we call a "site" already represents processed information; what we might term "type 1 information". This interpretive process extends further when establishing connections between sites. The "shared items" linking two locations aren't simply objects, but typological categories. The journey from individual artifacts (e.g. ceramic vessels) to established "types" involves complex analytical work. Here, "type 1 information" becomes the building material for a higher level of abstraction: "type 2 information".

<u>An illustrative case from Italian Prehistory</u>

The Bronze Age archaeology of Italy provides a compelling example of this process in action. In her influential work, Emma Blake (2014) relies on typologies established in the *Prähistorische Bronzefunde* (PBF) series to map connections between communities. PBF organizes bronze objects from a given country into typologies. Although it is a German project, many Italian scholars contributed, resulting in several volumes in which Italian bronze objects are classified into types and dated according to a relative chronological sequence based on their occurrence in different types and contexts. These typologies, carefully constructed classifications of bronze objects, serve as the author's network data, allowing her to identify specific clusters in the Late Bronze Age network, which she interprets as precursors to the ethnic groups that emerged in later phases.

The creation of such typological frameworks is intellectually demanding and time-consuming, traditionally appearing in comprehensive monographs rather than journal articles. Looking ahead, emerging technologies offer promising pathways. Machine learning approaches to artifact classification (Cardarelli 2024) may revolutionize how we generate and share this essential "type 2 information", a critical but currently undervalued component of archaeological knowledge production. These developments could fundamentally reshape not just how we analyse archaeological data, but how we publish and evaluate archaeological research.

Reimagining Metrics for Archaeological Innovation

What Drives Archaeological Discovery?

Innovation, elusive to define and challenging to quantify (Nadal et al. 2020), manifests in archaeology through three interconnected dimensions: conceptual breakthroughs, methodological advances, and new empirical foundations. These elements form a dynamic ecosystem: fresh theoretical frameworks inspire novel methodologies and demand previously unexplored data; innovative techniques unlock access to untapped evidence and catalyse conceptual evolution; and newly discovered data can fundamentally challenge established paradigms.

The Unique Epistemic Landscape of Archaeological Data

When we consider the distinctive nature of archaeological data discussed earlier, their production emerges as fundamentally important to the discipline for several compelling reasons. First, archaeology operates under conditions of unrepeatable observation, we cannot simply reconstruct the excavation that yielded our initial dataset. Second, identical methodological approaches can produce dramatically different results, also when applied to contemporaneous sites within the same region ("same experiment can give different results"). Third, as Wallach (2019) sharply observes, "inferences from absence [in archaeology] have an epistemic standing that is comparable to other empirical inferences".

The unique nature of archaeological data necessitates scientific outputs that are deeply focused on specific contexts and research questions, both chronologically and geographically. These highly specialized studies form the essential building blocks for broader interpretative frameworks and synthetic narratives.

Some archaeological subfields, such as Palaeolithic research, often address themes with wider appeal and broader relevance, making them more likely to be accepted by high-impact journals. In contrast, Italian research tradition tends to adopt for later epochs, like the Bronze Age, a micro-scale and regional approach, resulting in detailed and focused investigations that are less frequently published in major journals. This discrepancy reflects the challenges faced by specialists working in complex, heterogeneous fields where the value of meticulous, context-dependent research may be underappreciated in the competitive landscape of academic publishing.

Recognizing this dynamic is crucial to understanding the epistemic foundations of the production of prehistoric archaeological knowledge, and the need for diverse publication formats that accommodate both specialized case studies and synthetic overviews.

Archaeological interpretation advances toward stability only when we achieve a "critical mass" of classified specimens, that threshold where additional examples no longer necessitate the creation of new typological categories, allowing us to construct increasingly robust hypotheses.

Toward Archaeological-Specific Metrics: Lessons Learned from Italian Prehistoric Archaeology.

These considerations as shown by the Italian case study demand a fundamental recalibration of how we measure scholarly impact in prehistoric archaeology (and perhaps archaeology in general):
Elevating Monographs: Indexing systems must incorporate monographs, particularly those published in series with rigorous quality protocols comparable to peer-reviewed journals. Works often dismissed as mere "catalogues" should be recognized for their crucial role in generating type 1 and type 2 information. Citation metrics should track the most widely used datasets, identifying truly groundbreaking contributions to the field.
Recognizing Regional Specialization: Indexes must account for journals focused on specific regions and chronological periods, publications that provide the deep contextual understanding essential for interpreting the fragmented evidence that constitutes archaeology's empirical foundation.
Embracing Digital Transformation: As a discipline, we should champion the creation of digital repositories for type 1 and type 2 datasets, indexed with the same rigor as traditional monographs. These collections could pair with specialized journals publishing papers that examine metadata structures and data curation methodologies.
Valuing Depth Over Volume: To effectively utilize metrics based on these reimagined indexes, we must consider the scope and magnitude of individual contributions rather than simply counting outputs. While the logic of hyperproductivity has faced widespread criticism across academia, archaeology faces a particular truth: in our field, it's "Time" that ultimately counts.

Conclusions

This comprehensive examination of bibliographic indexing and research metrics, as it is shown by the case study, reveals a scientific publishing ecosystem reveals an ongoing conflict between its foundational mission, advancing human knowledge, and the structural forces that increasingly govern it. The evolution from early indexing systems like *Index Medicus* to contemporary platforms such as Web of Science, Scopus, and Google Scholar reflects not merely technological progress, but a fundamental transformation in how scientific value is conceptualized, measured, and distributed.
This analysis demonstrates that the apparent objectivity of citation-based metrics masks significant epistemological challenges. The comparative study of archaeological journals across multiple indexes reveals how varying inclusion criteria and algorithmic approaches produce dramatically different representations of scholarly output within a single discipline. When only one journal appears in all five major indexes examined, and when highly cited works remain invisible to "trusted" databases, we must question whether current evaluation systems truly capture scientific significance or merely reflect the biases embedded in their selection mechanisms.
The disciplinary analysis of archaeology illuminates broader issues affecting Humanities and Social sciences. The underrepresentation of regionally focused journals in international indexes, the inadequate handling of monographs, archaeology's primary vehicle for publishing empirical data, and the mismatch between citation patterns and actual scholarly influence all point to fundamental limitations in applying standardized metrics across diverse knowledge domains. The concept of "type 1" and "type 2"

information reveals how disciplines with unique epistemic foundations, as it is archaeological research, require correspondingly specialized approaches to evaluation.

Perhaps most concerning is the evidence of how metrics have begun to reshape scientific practice itself rather than merely measuring it. The documented decline in disruptive research and the rise of unethical behaviours represents symptoms of a system that has prioritized measurability and quantity over meaningfulness and that require, above all, a cultural change. The "publish or perish" mentality, originally conceived as "publish lest the knowledge should perish with you", has been inverted into a productivity imperative that may actually impede the preservation and advancement of knowledge.
The peer review process, positioned as science's primary quality control mechanism, emerges from this analysis as a system operating more on faith than evidence. The documented inconsistencies, biases, and limited effectiveness of traditional peer review, combined with the promising developments in open peer review and post-publication evaluation platforms, suggest that quality assurance in scientific publishing requires fundamental reconceptualization rather than mere refinement.
Yet this analysis also identifies pathways toward more equitable and effective evaluation systems. The principles articulated in DORA and the Leiden Manifesto provide frameworks for contextualizing metrics rather than abolishing them. The emergence of overlay journals, preprint repositories, and platforms like PubPeer demonstrates the scientific community's capacity for self-correction and innovation.

For archaeology specifically, and for Humanities and Social sciences more broadly, the path forward requires embracing disciplinary specificity while maintaining scholarly rigor. This means developing indexing systems that account for monographs and regional publications, creating metrics that value depth and contextual significance over volume, and fostering digital repositories that make specialized datasets accessible for citation and reuse. The field's unique relationship with unrepeatable data and its dependence on accumulated evidence demands evaluation criteria that recognize the long-term, collaborative nature of knowledge building.

Ultimately, this study argues for a more nuanced, pluralistic approach to scientific assessment, one that recognizes that meaningful evaluation cannot be achieved through metrics alone, but requires human judgment informed by disciplinary expertise and contextual understanding. The goal should not be to eliminate quantitative measures, but to deploy them more thoughtfully within evaluation frameworks that tribute the diversity of scientific practice and the complexity of intellectual contribution.
The transformation of scientific publishing and evaluation will require sustained effort from all stakeholders: researchers who must resist the temptation of gaming metrics while advocating for fair assessment; editors and publishers who must balance commercial viability with scholarly integrity; institutions that must develop promotion and tenure criteria reflecting the full spectrum of academic contributions; and funding agencies that must support both high-risk, innovative research and the foundational work that enables future breakthroughs.

As the Italian case study demonstrates, national evaluation systems like ANVUR can play crucial roles in recognizing regional scholarship and disciplinary specificity, but they must remain vigilant against reproducing the limitations of international indexes at a local level. The challenge lies in maintaining scientific standards while ensuring that evaluation systems serve science rather than constraining it.

In closing, this analysis suggests that the current crisis in scientific publishing and evaluation may ultimately prove catalytic, forcing a necessary reckoning with systems that have outlived their utility. The emergence of alternative models, from overlay journals to post-publication peer review, indicates that the scientific community possesses both the will and the creativity to develop more equitable and effective approaches to knowledge dissemination and assessment. The question is not whether change will come, but whether it will be guided by principles that truly serve the advancement of human understanding.

The path forward requires acknowledging that science, at its best, remains fundamentally a human endeavour driven by curiosity, collaboration, and the patient accumulation of knowledge across generations. Any evaluation system worthy of this enterprise must reflect these values while adapting to the realities of contemporary academic life. In archaeology, as in all disciplines, the ultimate measure of our success should not be the efficiency with which we produce publications, but the depth and durability of the understanding we create.

References

Academia, 8 Feb 2025. Accessed 8 Feb 2025.
https://www.academia.edu/about

Aczel B., Szaszi B. & Holcombe A. O. 2021. A billion-dollar donation: estimating the cost of researchers' time spent on peer review. *Research Integrity and Peer Review* 6: 14. https://doi.org/10.1186/s41073-021-00118-2.

Allum N. 2024. How common is academic plagiarism? - Impact of Social Sciences. *Impact of Social Sciences - Maximizing the impact of academic research*. https://blogs.lse.ac.uk/impactofsocialsciences/2024/02/08/how-common-is-academic-plagiarism/.

Altmetric, 20 Feb 2025. Accessed 20 Feb 2025.
https://www.altmetric.com/

Amaral O.B. 2018. All publishers are predatory - some are bigger than others. *Anais da Academia Brasileira de Ciências* 90: 1643-1647. https://doi.org/10.1590/0001-3765201820170959.

AMiner, 15 Feb 2025. Accessed 15 Feb 2025.
https://www.aminer.cn/

ANVUR 30 June 2025. Accessed 30 Jun 2025.
https://www.anvur.it/it/ricerca/riviste

arXiv, 20 March 2025. Accessed Mar 2025.
https://arxiv.org/

Bartneck C. 2017. Reviewers' scores do not predict impact: bibliometric analysis of the proceedings of the human–robot interaction conference. *Scientometrics* 110: 179-194. https://doi.org/10.1007/s11192-016-2176-y.

BEALL J. 2010. "Predatory" open-access scholarly publishers. *The Charleston Advisor* 11: 10-17.

—. 2012. Predatory publishers are corrupting open access. *Nature* 489: 179–179. https://doi.org/10.1038/489179a.

—. 2017. What I learned from predatory publishers. *Biochemia Medica* 27: 273–78. https://doi.org/10.11613/BM.2017.029.

BEEL J. & GIPP B. 2009. Google Scholar's Ranking Algorithm : An Introductory Overview., in *Proceedings of the 12th International Conference on Scientometrics and Informetrics (ISSI'09)*, 1: 230-241. Rio de Janeiro (Brazil): International Society for Scientometrics and Informetrics.

BERGSTROM C. 2007. Eigenfactor: Measuring the value and prestige of scholarly journals. *College & Research Libraries News* 68: 314-316. https://doi.org/10.5860/crln.68.5.7804.

BJÖRK B., ROOS A. & LAURI M 2008. Global annual volume of peer reviewed scholarly articles and the share available via different open access options. CHAN L. & MORNATI S. (eds.), *Open Scholarship: Authority, Community, and Sustainability in the Age of Web 2.0 - Proceedings of the 12th International Conference on Electronic Publishing held in Toronto, Canada 25-27 June 2008*. Toronto: pp. 178-186.

BLAKE, E. 2014. *Social networks and regional identity in Bronze Age Italy*. New York, NY: Cambridge University Press.

BREZIS E.S. & BIRUKOU A. 2020. Arbitrariness in the peer review process. *Scientometrics* 123: 393-411. https://doi.org/10.1007/s11192-020-03348-1.

BROOKER R. & ALLUM N. 2024. Investigating the links between questionable research practices, scientific norms and organisational culture. *Research Integrity and Peer Review* 9: 12. https://doi.org/10.1186/s41073-024-00151-x.

BURANYI S. 2017. Is the staggeringly profitable business of scientific publishing bad for science? *The Guardian*, June 27, sec. Science. https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science.

Cabell, 18 Mar 2025. Accessed 18 Mar 2025. https://cabells.com/

CAMPANARIO J.M. 2009. Rejecting and resisting Nobel class discoveries: accounts by Nobel Laureates. *Scientometrics* 81: 549-565. https://doi.org/10.1007/s11192-008-2141-5.

CANDAL-PEDREIRA C., ROSS J.S., RUANO-RAVINA A., EGILMAN D.S., FERNÁNDEZ E. & PÉREZ-RÍOS M. 2022. Retracted papers originating from paper mills: cross sectional study. *BMJ*, e071517. https://doi.org/10.1136/bmj-2022-071517.

CARDARELLI. *Morphological variability and standardisation of vessel shapes in the 2nd and first half of the first millennium BC in continental Italy* (Adrias 18). Edipuglia.

CARLSON P.G. 1928. Periodicals and their indexing. *Peabody Journal of Education* 6: 30-32. https://doi.org/10.1080/01619562809534850.

Clarivate History, 10 Feb 2025. Accessed 10 Feb 2025. https://clarivate.com/academia-government/the-institute-for-scientific-information/history/

Clarivate Editorial Selection Process, 10 Feb 2025. Accessed 10 Feb 2025. https://clarivate.com/academia-government/scientific-and-academic-

research/research-discovery-and-referencing/web-of-science/web-of-science-core-collection/editorial-selection-process/
COBEY K.D., LALU M.M., SKIDMORE B., AHMADZAI N., GRUDNIEWICZ A. & MOHER D. 2018. What is a predatory journal? A scoping review. *F1000Research* 7: 1001. https://doi.org/10.12688/f1000research.15256.2.
COCI, 20 Feb 2025. Accessed 20 Feb 2025. https://opencitations.net/
COPE paper mill, 18 Mar 2025. Accessed 18 Mar 2025. https://publicationethics.org/topic-discussions/paper-mills
CrossRef, 13 Feb 2025. Accessed 13 Feb 2025. https://www.crossref.org/
CrossRef GIT, 13 Feb 2025. Accessed 13 Feb 2025. https://gitlab.com/crossref/retraction-watch-data
DE GUIO A. 1988. Unità archeostratigrafiche cone unità operazionali: verso le archeologie possibili degli anni '90. *Quaderni di Archeologia Stratigrafica dell'Italia Settentrionale* 1: 9-22.
Dean J.W., Flower P.B.S. 1985. Peer review at work. *British Medical Journal* 29: 1555-1561.
Dimension 13 Feb 2025. Accessed 13 Feb 2025. https://www.dimensions.ai/
DI RENZONI A. 2025. Indexing Science. Archeologia e Calcolatori 36.1: 509-513.
DONY C., RASKINET M., RENAVILLE F., SIMON S. & THIRION P. 2020. How reliable and useful is Cabell's Blacklist ? A data-driven analysis. *LIBER Quarterly* 30: 1. https://doi.org/10.18352/lq.10339.
Dora, 10 Mar 2025. Accessed 10 Mar 2025. https://sfdora.org/
EGGHE L. 2006. Theory and practise of the g-index. *Scientometrics* 69: 131-152. https://doi.org/10.1007/s11192-006-0144-7.
FRANGIPANE M. 2023. Scienza e Scienza. Superamento dei confini, Slow Science e libertà della ricerca. Presented at the Inaugurazione Anno Accademico dell'Accademia dei Lincei, Anno 2023-24. https://www.lincei.it/sites/default/files/documenti/20231110_M_Frangipane_A5.pdf.
FUNK R.J. & OWEN-SMITH J. 2017. A Dynamic Network Measure of Technological Change. *Management Science* 63: 791-817. https://doi.org/10.1287/mnsc.2015.2366.
FYFE A., MOXHAM N., MCDOUGALL-WATERS J. & MØRK RØSTVIK C. 2022. *A History of Scientific Journals*. UCL Press. https://discovery.ucl.ac.uk/id/eprint/10156072/. https://doi.org/10.14324/111.9781800082328.
GFTR, 10 March 2025. Accessed 10 March 2025. https://www.getfulltextresearch.com/
GODLEE F. & JEFFERSON T. (eds.) 1999. *Peer review in health sciences* (Medical Research). London: BMJ Books.
GREENBERG S.J. & GALLAGHER P. E. 2009. The great contribution: Index Medicus, Index-Catalogue, and IndexCat. *Journal of the Medical Library Association : JMLA* 97: 108-113. https://doi.org/10.3163/1536-5050.97.2.007.
GRUDNIEWICZ A. et al. 2019. Predatory journals: no definition, no defence. *Nature* 576:

210–12. https://doi.org/10.1038/d41586-019-03759-y.
HANSON M.A., BARREIRO P. G., CROSETTO P. & BROCKINGTON D. 2024. The strain on scientific publishing. *Quantitative Science Studies* 5: 823-843.
https://doi.org/10.1162/qss_a_00327.
Harzing's Publish or Perish, 10 March 2025.
https://harzing.com/resources/publish-or-perish
HERBERT R. 2020. Accept Me, Accept Me Not: What Do Journal Acceptance Rates Really Mean? [ICSR Perspectives]. SSRN. https://www.ssrn.com/abstract=3526365.
https://doi.org/10.2139/ssrn.3526365.
HICKS D., WOUTERS P., WALTMAN L., DE RIJCKE S. & RAFOLS I. 2015. Bibliometrics: The Leiden Manifesto for research metrics. *Nature* 520: 429–31.
https://doi.org/10.1038/520429a.
HIRSCH J.E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences* 102: 16569-16572.
https://doi.org/10.1073/pnas.0507655102.
IBRAHIM H., LIU F., ZAKI Y. & RAHWAN T. 2025. Citation manipulation through citation mills and pre-print servers. *Scientific Reports* 15: 5480. https://doi.org/10.1038/s41598-025-88709-7.
JINHA A.E. 2010. Article 50 million: an estimate of the number of scholarly articles in existence. *Learned Publishing* 23: 258-263. https://doi.org/10.1087/20100308.
KIM K. & CHUNG Y. 2018. Overview of journal metrics. *Science Editing* 5: 16–20.
https://doi.org/10.6087/kcse.112.
KING D.W., TENOPIR C., CHOEMPRAYONG S. & WU L. 2009. Scholarly journal information-seeking and reading patterns of faculty at five US universities. *Learned Publishing* 22: 126–44. https://doi.org/10.1087/2009208.
KULCZYCKI E. et al. 2018. Publication patterns in the social sciences and humanities: evidence from eight European countries. *Scientometrics* 116: 463–86.
https://doi.org/10.1007/s11192-018-2711-0.
LARIVIÈRE V. & SUGIMOTO C.R. 2019. The Journal Impact Factor: A Brief History, Critique, and Discussion of Adverse Effects, in Glänzel W., Moed H.F., Schmoch U. & Thelwall M. (eds.) *Springer Handbook of Science and Technology Indicators*: 3–24 (Springer Handbooks). Cham: Springer International Publishing.
http://link.springer.com/10.1007/978-3-030-02511-3_1. https://doi.org/10.1007/978-3-030-02511-3_1.
LARIVIÈRE V., HAUSTEIN S. & MONGEON P. 2015. The Oligopoly of Academic Publishers in the Digital Era. *PLOS ONE* 10: e0127502.
https://doi.org/10.1371/journal.pone.0127502.
LAWRENCE, A. 2024. Harvard's Claudine Gay was ousted for 'plagiarism'. How serious was it really? *The Guardian*, January 6, sec. Education.
https://www.theguardian.com/education/2024/jan/06/harvard-claudine-gay-plagiarism.
Libkey, 10 Mar 2025. Accessed 10 Mar 2025
https://libkey.io/
LOCK S. 1985. *A difficult balance: editorial peer review in medicine* (The Rock Carling

Fellowship 1985). Philadelphia: ISI Pr.
MABE M. 2003. The growth and number of journals. *Serials: The Journal for the Serials Community* 16: 191-197. https://doi.org/10.1629/16191.
MARCUS A. & ORANSKY I. 2011. The paper is not sacred. *Nature* 480: 449–50. https://doi.org/10.1038/480449a.
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: 871–906. https://doi.org/10.1007/s11192-020-03690-4.
Microsoft Academic, 10 Feb 2025. Accessed 10 Feb 2025.
https://www.microsoft.com/en-us/research/articles/microsoft-academic-to-expand-horizons-with-community-driven-approach/
MOED H.F. 2010. Measuring contextual citation impact of scientific journals. *Journal of Informetrics* 4: 265-277. https://doi.org/10.1016/j.joi.2010.01.002.
MOSKOVKIN V. 2024. Tracing the origins of 'publish or perish' - Impact of Social Sciences. *Impact of Social Sciences - Maximizing the impact of academic research*.
https://blogs.lse.ac.uk/impactofsocialsciences/2024/07/15/tracing-the-origins-of-publish-or-perish/.
NANDAL N., AARUSHI K. & MEENAKSHI D. 2020. Measuring Innovation: Challenges and Best Practices. *International Journal of Advanced Science and Technology* 29: 1275–85.
NEYLON C. 2015. Researchers are not 'hoodwinked' victims. All choose to play the publishing game and some can choose to change it. - Impact of Social Sciences. *Impact of Social Sciences - Maximizing the impact of academic research*.
https://blogs.lse.ac.uk/impactofsocialsciences/2015/09/11/researcher-as-victim-researcher-as-predator/.
NICHOLSON J.M., MORDAUNT M., LOPEZ P., UPPALA A., ROSATI D., RODRIGUES N. P., GRABITZ P. & RIFE S. C. 2021. scite: A smart citation index that displays the context of citations and classifies their intent using deep learning. *Quantitative Science Studies* 2: 882-898. https://doi.org/10.1162/qss_a_00146.
OH K. 2023. Get Full Text Research (GetFTR): can it be a good tool for researchers? *Science Editing* 10: 186–89. https://doi.org/10.6087/kcse.311.
OLIJHOEK T. & TENNANT J. 2018. The "problem" of predatory publishing remains a relatively small one and should not be allowed to defame open access - Impact of Social Sciences. *Impact of Social Sciences - Maximizing the impact of academic research*.
https://blogs.lse.ac.uk/impactofsocialsciences/2018/09/25/the-problem-of-predatory-publishing-remains-a-relatively-small-one-and-should-not-be-allowed-to-defame-open-access/.
OLIVE R., TOWNSEND S. & PHILLIPS M.G. 2023. 'Not everything that can be counted counts, and not everything that counts can be counted': Searching for the value of metrics and altmetrics in sociology of sport journals. *International Review for the Sociology of Sport* 58: 431-454. https://doi.org/10.1177/10126902221107467.
OpenAI, 18 Mar 2025. Accessed 18 Mar 2025.
https://openai.com/research/
OpenAlex, 13 Feb 2025. Accessed 13 Feb 2025.
https://openalex.org/

ORTEGA J.L. 2022. Classification and analysis of PubPeer comments: How a web journal club is used. *Journal of the Association for Information Science and Technology* 73: 655–70. https://doi.org/10.1002/asi.24568.

OZONOFF D. 2024. As the world turns: scientific publishing in the digital era. *Environmental Health* 23: 24, s12940-024-01063–65. https://doi.org/10.1186/s12940-024-01063-5.

PARK M., LEAHEY E. & FUNK R. J. 2023. Papers and patents are becoming less disruptive over time. *Nature* 613: 138-144. https://doi.org/10.1038/s41586-022-05543-x.

Peer Community, 18 Mar 2025. Accessed 18 Mar 2025. https://peercommunityin.org

Peer Community Archaeology, 18 Mar 2025. Accessed 18 Mar 2025. https://archaeo.peercommunityin.org/

PEERE, 14 Mar 2025. Accessed 14 Mar 2025. https://www.peere.org/

PETERS D.P. & CECI S.J. 1982. Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences* 5: 187-195. https://doi.org/10.1017/S0140525X00011183.

PHILLIPS M.G. 2020. Sizing up Sport History Journals: Metrics, Sport Humanities, and History. *The International Journal of the History of Sport* 37: 692-704. https://doi.org/10.1080/09523367.2020.1796652.

PubPeer, 20 Mar 2025. Accessed 20 Mar 2025. https://pubpeer.com/static/about

PUPOVAC V. & FANELLI D. 2015. Scientists Admitting to Plagiarism: A Meta-analysis of Surveys. *Science and Engineering Ethics* 21: 1331-1352. https://doi.org/10.1007/s11948-014-9600-6.

RAGONE A., MIRYLENKA K., CASATI F. & MARCHESE M. 2013. On peer review in computer science: analysis of its effectiveness and suggestions for improvement. *Scientometrics* 97: 317-356. https://doi.org/10.1007/s11192-013-1002-z.

ReserachGate, 8 Feb 2025. Accessed 8 Feb 2025. https://help.researchgate.net/hc/en-us

Retraction Database, 18 Mar 2025. Accessed 18 Mar 2025. https://retractiondatabase.org/RetractionSearch.aspx?

Scopus Metrics, 10 Feb 2025. Accessed 10 Feb 2025. https://www.elsevier.com/products/scopus/metrics

Scopus Selection Criteria, 10 Feb 2025. Accessed 10 Feb 2025 https://www.elsevier.com/products/scopus/content/content-policy-and-selection

SEGLEN P.O. 1997. Citations and journal impact factors: questionable indicators of research quality. *Allergy* 52: 1050-1056. https://doi.org/10.1111/j.1398-9995.1997.tb00175.x.

SEPPELT R., BECKMANN M., VÁCLAVÍK T. & VOLK M. 2018. The Art of Scientific Performance. *Trends in Ecology & Evolution* 33: 805–809. https://doi.org/10.1016/j.tree.2018.08.003.

Scilit, 15 Feb 2025. Accessed 15 Feb 2025. https://www.scilit.com/

Semantic Scholar, 20 Feb 2025. Accessed 20 Feb 2025.
https://www.semanticscholar.org/
SILER K. 2020. There is no black and white definition of predatory publishing - Impact of Social Sciences. *Impact of Social Sciences - Maximizing the impact of academic research*. https://blogs.lse.ac.uk/impactofsocialsciences/2020/05/13/there-is-no-black-and-white-definition-of-predatory-publishing/.
SMITH R. 1999. Opening up BMJ peer review. *BMJ (Clinical research ed.)* 318: 4-5. https://doi.org/10.1136/bmj.318.7175.4.
SMITH R. 2006. Peer Review: A Flawed Process at the Heart of Science and Journals. *Journal of the Royal Society of Medicine* 99: 178-182. https://doi.org/10.1177/014107680609900414.
Sokal affaire, 15 Mar 2025. Accessed 15 Mar 2025.
https://en.wikipedia.org/wiki/Sokal_affair
SQUAZZONI F., BREZIS E. & MARUŠIĆ A. 2017. Scientometrics of peer review. *Scientometrics* 113: 501-502. https://doi.org/10.1007/s11192-017-2518-4.
STRIELKOWSKI W. 2018. Predatory Publishing: What Are the Alternatives to Beall's List? *The American Journal of Medicine* 131: 333–34. https://doi.org/10.1016/j.amjmed.2017.10.054.
ŠUPAK SMOLČIĆ V. 2013. Salami publication: definitions and examples. *Biochemia Medica*, 237–41. https://doi.org/10.11613/BM.2013.030.
SUPAK SMOLCIC V. & BILIC-ZULLE L. 2013. How do we handle self-plagiarism in submitted manuscripts? *Biochemia Medica*, 150-153. https://doi.org/10.11613/BM.2013.019.
SWAUGER S. 2017. Open access, power, and privilege: A response to "What I learned from predatory publishing". *College & Research Libraries News* 78: 603. https://doi.org/10.5860/crln.78.11.603.
TENOPIR C. & KING D.W. 2014. The growth of journals publishing, in COPE B. & Phillips A. (eds.) *The Future of the Academic Journal (Second Edition)*: 159-78. Chandos Publishing.
https://www.sciencedirect.com/science/article/pii/B9781843347835500069.
https://doi.org/10.1533/9781780634647.159.
The conceptual penis, 15 Mar 2025. Accessed 15 Mar 2025.
https://www.tandfonline.com/doi/full/10.1080/23311886.2017.1336861
The Lens, 13 Feb 2025. Accessed 13 Feb 2025.
https://www.lens.org/
UKRI, 10 Mar 2025. Accessed 10 Mar 2025.
https://www.ukri.org/publications/review-of-metrics-in-research-assessment-and-management/
WAGER, E., SINGHVI S. & KLEINERT S. 2015. Too much of a good thing? An observational study of prolific authors. *PeerJ* 3: e1154. https://doi.org/10.7717/peerj.1154.
WALLACH E. 2019. Inference from absence: the case of archaeology. *Palgrave Communications* 5: 94. https://doi.org/10.1057/s41599-019-0307-9.
WELLISCH H. H. 1983. 'Index': the word, its history, meanings and usages. *The Indexer* 13: 147-51. https://doi.org/10.3828/indexer.1983.13.3.2.

WOLFRAM D., WANG P., HEMBREE A. & PARK H. 2020. Open peer review: promoting transparency in open science. *Scientometrics* 125: 1033-1051. https://doi.org/10.1007/s11192-020-03488-4.
XIE, Y., WANG K. & KONG Y. 2021. Prevalence of Research Misconduct and Questionable Research Practices: A Systematic Review and Meta-Analysis. *Science and Engineering Ethics* 27: 41. https://doi.org/10.1007/s11948-021-00314-9.
YOUNG N.S., IOANNIDIS J.P.A. & AL-UBAYDLI O. 2008. Why Current Publication Practices May Distort Science. *PLoS Medicine* 5: e201. https://doi.org/10.1371/journal.pmed.0050201.

Appendix

Lexicon

– **Abstracting database**: A database that provides structured summaries (abstracts) of academic articles, reports, or other documents. These summaries allow users to assess the relevance of a document without reading the full text.
– **Indexing database**: An organized collection of documents, assigning subject terms, keywords, and other metadata to facilitate retrieval. Indexing improves discoverability by enabling users to search for documents based on structured criteria rather than full-text searches.
– **Aggregators**: Platforms that collect and distribute scholarly content from multiple publishers, databases, or academic institutions. Aggregators provide centralized access to a wide range of scientific products, often through subscriptions or institutional licenses.
– **Citation index**: A specialized type of bibliographic database that tracks citations between published works.
– **Abstracting and indexing service (A&I service)**: A commercial or institutional product that provides both abstracting and indexing functions. These services are offered by publishers, academic institutions, or specialized information providers.

Indexes & Other Services

Main indexes, databases & services

The number of databases and A&I services has grown rapidly with the spread of the internet, aiming to link scientific data and publications and, importantly, to manage citations among papers. Among the major indexing services, such as Scopus, Web of Science, and Google Scholar (described in the text), there are also other notable platforms worth mentioning.

Microsoft Academic was a free, AI-driven academic search engine developed by Microsoft Research. It provided access to a vast collection of scholarly publications, authors, institutions, and research topics, leveraging natural language processing (NLP) to extract and analyze metadata. It played a key role in academic discovery until its shutdown in December 2021, after which its data (Microsoft Academic Graph, containing billions of entities, including papers, authors, journals, citations, and more, organized as a relational graph) contributed to the development of OpenAlex and other research indexing projects (Microsoft Academic, 10 Feb 2025).

OpenAlex is a free and open catalogue of global scholarly research aimed at indexing academic outputs. It allows users to search using various metadata fields. Works are ranked solely by citations. Its primary data source is Microsoft Academic Graph. Data can be accessed via a free API (OpenAlex, 13 Feb 2025).

CrossRef is a nonprofit organization providing open digital infrastructure for the scholarly research community. It is the largest Digital Object Identifier (DOI) Registration Agency of the International DOI Foundation. Rather than hosting full-text scientific content, CrossRef facilitates links between distributed content on external sites through open metadata and persistent identifiers. It manages citation linking and offers services such as "Similarity Check", a plagiarism detection tool. (CrossRef, 13 Feb 2025).

Dimensions is a comprehensive research data infrastructure that enables users to explore connections across various scholarly outputs. It integrates a wide range of linked data, including grants, publications and datasets. Publications are ranked based on: 1) total citations received from any publication type; 2) Field Citation Ratio; 3) recent citations, the number of citations received in the past two years; 4) Altmetric Attention Score. Publications and datasets can be browsed with a free account, while some features require a subscription (Dimension 13 Feb 2025).

AMiner is a free online service for indexing, searching, and mining scientific data. Users can browse content using metadata filters. Unlike other platforms, AMiner does not develop its own research metrics (AMiner, 15 Feb 2025).

The Lens aggregates metadata and full-text content, integrating scholarly works, patents, and biological sequences with management tools (The Lens, 13 Feb 2025).

Scilit is a multidisciplinary, free scholarly database and aggregator that indexes scientific literature by harvesting up-to-date metadata from sources such as CrossRef, PubMed, and other repositories. It offers users tools to rank and evaluate editors, journals, preprint servers, and institutions using a variety of metrics, facilitating discovery and assessment of scientific output (Scilit, 15 Feb 2025).

Semantic Scholar is an AI-powered research discovery platform that enhances academic search through natural language processing (NLP) and machine learning. It extracts key insights, summarizes findings, and identifies influential papers across millions of publications. Semantic Scholar prioritizes context and impact analysis rather than simple keyword matching. The platform covers a wide range of disciplines, with a strong emphasis on computer science, biomedical research, and AI-related fields (Semantic Scholar, 20 Feb 2025).

COCI (OpenCitations Index) is an open and freely accessible index of citation data derived from CrossRef. It provides structured bibliographic and citation metadata. COCI follows open data principles, allowing unrestricted reuse of its content. Users can access and query COCI data via SPARQL endpoints, REST APIs, and bulk downloads, making it a valuable resource for bibliometric analysis and scholarly network exploration (COCI, 20 Feb 2025).

Academic Social network

Among the major indexing services such as Scopus, Web of Science, and Microsoft Academic Graph, platforms like ResearchGate and Academia.edu, although not formal indexing databases, contribute significantly to the dissemination of research and, in some cases, provide their own forms of author and paper rankings.

ResearchGate is a social networking platform designed for researchers and scientists to share their work, collaborate, and engage with the global academic community. It allows users to upload publications, ask and answer research-related questions, and track citation metrics. It does not function as a formal indexing service. Instead, it serves as a repository where researchers can disseminate their work and connect with peers, complementing traditional academic databases and indexing platforms. ResearchGate generates its own citation index based on full-text documents found by its web crawler and those uploaded by users. Citations, along with other indicators, are used to calculate the Research Interest Score.
In addition to RIS, ResearchGate uses two versions of the h-index (discussed below) to rank authors, one that includes self-citations and another that excludes them. The h-index is calculated solely based on the publications listed in users' profiles (ReserachGate, 8 Feb 2025).
Academia.edu is an online platform that enables researchers to share their publications, follow topics of interest, and engage with the academic community. It serves as a repository where scholars can upload their work, discover relevant research, and track readership metrics. While Academia.edu enhances visibility and networking opportunities, it is not a formal indexing service. Unlike traditional academic databases, it operates on a freemium model, offering additional analytics and promotional features to paying users. Academia.edu provides engagement metrics such as profile views, document count and downloads, and mentions. However, it does not have a proper scoring system (Academia, 8 Feb 2025).

Other services

In addition to indexing databases and academic social networks, there are several specialized tools and services that support scholarly publishing by facilitating citation analysis, seamless access to full-text articles, and monitoring publication integrity. Notable examples include Publish or Perish, Get Full Text Research (GetFTR), LibKey, and the Retraction Watch database.

Harzing's Publish or Perish is a software program that retrieves and analyses academic citations from multiple data sources, providing a range of citation metrics such as the number of papers, total citations, and the h-index (Harzing's Publish or Perish, 10 March 2025).

Get Full Text Research (GFTR) is a service that provides researchers with direct access to online journal articles by leveraging existing access technologies, such as IP-based authentication and federated access. This allows users to eliminate the need for researchers to manually log into their institution's library system beforehand. Recently, GetFTR introduced a browser extension that researchers can install themselves. (Kwangil Oh 2023; GFTR, 10 March 2025)

LibKey is a tool designed to provide seamless full-text access for researchers. It provides users with smooth and reliable full text linking experience by leveraging artificial intelligence to select sources, generate PDF links, and interpret open access (OA) availability. It offers intelligent link classification and manages user authentication (Libkey, 10 Mar 2025).

Retraction Watch Database is a tool that tracks retracted articles, many of which are withdrawn due to unethical practices. Created by Ivan Oransky and Adam Marcus, it provides an updated record of retractions and, since its acquisition by Crossref in September 2023, is maintained with daily updates and made available as a downloadable CSV file via GitHub. (CrossRef GIT).

Metrics

Journal Metrics

The Journal Impact Factor (JIF) is perhaps the most widely recognized metric for evaluating journal performance, although its reliability has been increasingly questioned in recent years (Larivière, Sugimoto 2019). Introduced by Eugene Garfield in 1964, the JIF remains a key tool for ranking journals and is published annually in the Journal Citation Reports (JCR), a service currently managed by Clarivate Analytics. It is extensively used by academic communities worldwide.

The JIF is calculated each year by dividing the total number of citations received in the current year by articles published in a journal during the previous two years by the total number of citable items published in that journal over the same period. Citable items typically include research articles and reviews but exclude editorials, letters, and abstracts. Citations contributing to the JIF come from various document types, including papers and other scholarly works.

Related metrics include the 5-year Impact Factor, which considers citations over a five-year period rather than two, and the Immediacy Index, which measures citations received within the same year as publication.

An alternative metric, the Impact Factor without self-citations (IFwoSC), excludes citations that a journal receives from its own articles. A significant discrepancy between the standard JIF and the IFwoSC can lead to a journal's removal from the JCR list, as this serves to prevent citation manipulation.

The JIF varies widely across disciplines due to differences in citation practices and publication speed. Therefore, specific metrics are published for each field. The Median Impact Factor (MIF) represents the JIF of the journal positioned in the middle when all

journals in a discipline are ranked by JIF. The Aggregate Impact Factor (AIF) is the ratio between total citations received by all papers in a discipline's journals and the total number of citable papers published in that discipline over the previous two years.
Despite its popularity, the JIF has been criticized for favoring disciplines with rapid citation cycles and for potentially encouraging practices aimed at inflating citation counts rather than reflecting research quality.
Cited Half Life (CHL) & Citing Half Life. CHL is a metric that calculates the median age of the citations received by a journal during the JCR year. The age of a citation is determined by subtracting the publication year of the cited item from the publication year of the citing item. For example, if a journal has a CHL of 6, it means that half of the citations refer to items published more recently than 6 years ago, while the other half refers to older items. It focuses on the citations a journal makes to other works. The CHL provides insight into a journal's relationships with its peers, indicating which journals it cites most often and how far back those citations reach (Kim, Chang 2018: 17).
Cite Score is calculated using data from the Scopus database. It is an index similar to the JIF, and it represents the average number of citations received per published paper in a specific journal over the previous three years. Both the numerator and denominator include all document types, not just research articles (Scopus Metrics, 10 Feb 2025).
Source Normalized Impact per Paper (SNIP), developed at Leiden University, measures a journal's contextual citation impact by considering the characteristics of its defined subject field. Specifically, it accounts for the frequency with which authors cite other papers in their reference lists, the speed at which citation impact matures, and the extent to which the database used for the assessment covers the field's literature (Moed 2010). SNIP is calculated by dividing the "raw impact per paper" (RIP) by the "relative database citation potential" (RDCP). RIP is the number of citations received in the year of calculation by papers published in the previous three years in a specific journal, divided by the total number of papers published. RDCP is determined as follows: Consider the references of papers that cited articles from the journal in the year X, where the cited papers were published in the previous three years. Among these references, include only those published during the same 3-year period. Divide the total number of those references by the number of citing papers.
Only citations from journals in the Scopus database are included, while citations from outside the database are ignored. The RDCP is then normalized by dividing the DCP by the median DCP of the database (Kim, Chang 2018: 19).
The **Scimago Journal Ranking (SJR)** indicator, developed by a research group from the Consejo Superior de Investigaciones Científicas (CSIC), University of Granada, Extremadura, Carlos III (Madrid), and Alcalá de Henares, uses data from the Scopus database. Similar to Eigenfactor, the SJR indicator is based on eigenvector centrality from network theory, where the importance of a node (journal) is determined by its connections to other high-scoring nodes.
The SJR calculation proceeds as follows: 1) Assign an initial score to each journal; 2) in iterative steps, the prestige of journals is redistributed through citations; 3) the iteration process continues until the difference in prestige values between consecutive iterations is smaller than a specified minimum threshold (Kim, Chang 2018: 19).

Eigenfactor (JES). The core concept of the JES is that journals are considered influential when they are frequently cited by other influential journals. This follows a procedure similar to the PageRank algorithm used by Google. JES is based on the eigenvector centrality algorithm, and represents a simple model in which researchers follow citations as they move from one journal to another. The process is as follows: 1) randomly select a journal article; 2) randomly choose one of the citations from the article; 4) proceed to the cited work; 5) select a new citation from this article; 6) repeat the process continuously.

The frequency with which each journal is visited reflects its importance within the academic citation network. Researchers tend to read journals that are highly cited by other influential journals. This iterative ranking model assumes that a single citation from a high-quality journal may carry more weight than multiple citations from less influential journals. The importance of a journal is thus measured by the influence of the citing journal divided by the total number of citations appearing in that journal (Berstrom 2007).

The Eigenfactor score of a journal indicates the percentage of time that journal is visited within the citation network. For example, if a journal has an Eigenfactor of 3.0, it means that 3% of the time, a researcher would be directed to this journal through the citation network.

Eigenfactor Scores tend to overestimate larger journals: the more articles a journal has, the more frequently it is expected to be visited. However, larger journals are not necessarily the most prestigious. To account for this, the Article Influence Score (AIS) is used. This index measures the influence of journals by considering citations per article and is directly comparable to the Journal Impact Factor (JIF). The AIS is calculated by dividing the Eigenfactor Score of a journal by the number of articles published, normalized so that the average article in the Journal Citation Reports (JCR) has an AIS of 1 (Kim, Chang 2018: 19).

Powered by scite_, **Smart Citations**, using deep learning models, categorizes citations based on their context. This feature provides a deeper understanding of journals, highlighting not only supportive mentions but also those that dispute or challenge the findings (Nicholson et al. 2021).

Although the **Acceptance Rate (AR)** is not typically used as a primary metric for ranking journals, many scholars consider it a straightforward measure. Journals with lower ARs are often perceived as more "prestigious". However, it is important to note that several factors can influence this index: 1) different journals may calculate AR in various ways, such as how they treat resubmissions, whether items like invited papers, special issues, and book reviews are included or excluded, and the timeframe considered; 2) some journals allow editors to select which manuscripts are even sent to the editorial team, and calculate their AR only based on these selected manuscripts, which is often less than the total number of submitted papers; 3) the number of submitted manuscripts may not be accurately recorded by editors; 4) highly specialized journals typically have a lower acceptance rate.

Some publishers, such as Wiley, display AR on their journal websites, but in many cases, this information must be requested from the editors. A 2020 study (Herbert 2020) of data from 2,300 journals, mostly published by Elsevier, analysed this metric and found the following: 1) the average AR is around 32%; 2) larger journals tend to have lower AR than smaller ones (ranging from 10% to 60%); 3) high-impact journals have relatively low AR, although there is considerable variation (from 5% to 50%); 4) gold open-access journals generally have higher AR than other types of open-access journals; 5) no clear relationship was found between a journal's scope and its AR, though STEM disciplines typically have lower ARs than journals in medicine or life sciences.

3.2 – Articles and authors metrics

Citation number. The most direct way to evaluate an article's impact is through the number of citations it receives. However, this number can vary significantly depending on the databases used to count citations. This metric is influenced by various factors, with some of the most significant being the discipline to which the paper pertains and the language in which it is written.

The **Altmetrics Attention Score (AAS)**, calculated by the Altmetric company, estimates the attention an article receives from non-traditional sources. It is designed to help identify how much and what types of attention a research output has garnered from different sources of attention, such as policy documents, news, blogs, social media. The score for an article increases as more people mention it. Different types of mentions contribute different base amounts to the final score. For example, a newspaper article contributes more than a blog post, which, in turn, contributes more than a tweet (Altmetric, 20 Feb 2025).

The **Consolidating or Disruptive index (CD index)** evaluates the impact of a paper on the discipline it pertains to, considering two possible outcomes: 1) some contributions enhance existing knowledge, thus consolidating the status quo; 2) some contributions disrupt existing knowledge, rendering it obsolete, and driving science and technology in new directions. It measures this distinction based on the premise that if a paper is disruptive, the subsequent work citing it is less likely to also cite its predecessors. Conversely, if a paper is consolidating, subsequent work that cites it is more likely to also cite its predecessors. The CD index ranges from -1 (consolidating) to 1 (disruptive).

The index is typically evaluated over a 5-year period from the paper's publication, as studies have shown that annual citations for most papers reach their peak within this time frame (Funk, Owen-Smith 2017).

The Field Citation Ratio (FCR) developed by Dimension is an article-level metric that indicates the relative citation performance of a publication when compared to other articles published in the same year within its subject area. A value greater than 1 indicates that the article has received more citations than the average for other articles published in the same subject area and year (Dimension, 13 Feb 2025).

Author level metrics

The H-index, proposed by Hirsch in 2005 (Hirsch 2005), is an author-level metric calculated using all of an author's papers, arranged according to the total number of citations they have received. The H-index is defined as the number of publications for which the author has been cited by other authors at least that same number of times (e.g., an H-index of 10 means the researcher has published at least 10 papers, each cited at least 10 times). The H-index can also be applied to journals using the total number of citations. However, the H-index has two major issues: since it is based on the total number of citations for each paper, it increases monotonically over time, even without the publication of new papers; researchers with a small number of very influential papers may have low H-indices.

To address these limitations, the G-index was proposed (Egghe 2006). It is calculated by arranging articles in decreasing order of citations and finding the largest number such that the first "g" papers together have at least "g^2" citations. For example, a G-index of 10 means that the top 10 papers by an author have received at least 100 citations.

Another similar metric provided by Google Scholar is the i10-index, which is the total number of papers authored by a researcher that have been cited at least 10 times.

The **Research Gate Interest Score (RIS)** combines reads by unique ResearchGate members, recommendations, and citations to provide a measure of an author's research impact. It focuses on individual research items and researchers' interactions with them. When a ResearchGate member reads, recommends, or cites a research item, that item's RIS increases according to the following weighting system: Read: 0.05; Full-text read: 0.15; Recommendation: 0.25; Citation: 0.5. RIS excludes self-citations, author reads, reads by non-ResearchGate members, multiple reads and recommendations by the same researcher within a single week, as well as interactions from bots, crawlers, and other automated systems (ResearchGate, 8 Feb 2025).

Table 1 . Ranking of the archaeological journals as calculated in the "index comparison"

Journals	Score	WoS	Scopus	Dimen.	Open Alex	GS	ANVUR
Nr Tot		165	516	2430	4501	na	2314
Journal of Archaeological Science	75,0	7	8	3	9	3	A
Journal of Cultural Heritage	52,0	3	4	11	0	1	A
Journal of Archaeological Method and Theory	40,8	4	6	0	15	8	A(2019)
Journal of Archaeological Research	32,4	1	2	0	6	0	A
Journal of World Prehistory	25,2	2	14	0	5	0	A(2019)
Antiquity	21,6	13	0	7	0	7	A)
Archaeological and Anthropological Sciences	20,4	8	19	0	0	2	A(2017
Journal of Anthropological Archaeology	16,8	10	0	0	14	11	A
American Antiquity	14,4	6	18	0	0	15	-
Science and Technology of Archaeological Research	12,8	5	5	0	0	0	-

Journals	Score	WoS	Scopus	Dimen.	Open Alex	GS	ANVUR
Radiocarbon	12,4	0	1	0	0	10	A
Vegetation History and Archaeobotany	10,8	0	13	0	19	13	A
Quaternary Science Reviews	7,6	0	3	0	20	0	A
Archaeometry	7,2	0	0	10	0	14	A
Digital Applications in Archaeology and Cultural Heritage	7,2	0	12	0	0	12	A(2018)
International Journal of Heritage Studies	7,2	0	0	20	0	4	-
Journal of Field Archaeology	4,4	0	0	12	0	19	A
Journal of Paleolithic Archaeology	4,4	14	0	0	0	17	-
Man (Journal of the Royal Anthropological Institute)	4,0	0	0	1	0	0	-
PNAS	4,0	0	0	0	1	0	A
American Journal of Physical Anthropology	3,8	0	0	2	0	0	A
Annual Review of Anthropology	3,8	0	0	0	2	0	A
Science	3,6	0	0	0	3	0	A
Cambridge Archaeological Journal	3,6	19	0	18	0	20	A
Journal of the American Oriental Society	3,4	0	0	4	0	0	A
Nature Communications	3,4	0	0	0	4	0	A
Heritage Science	3,2	0	0	0	0	5	-
Journal of Forensic Sciences	3,2	0	0	5	0	0	-
Virtual Archaeology Review	3,2	18	16	0	0	0	-
World Archaeology	3,2	16	0	0	18	0	A
Journal of Archaeological Science: reports	3,0	0	0	0	0	6	-
The Journal of Roman Studies	3,0	0	0	6	0	0	A
Frontiers of Architectural Research	2,8	0	7	0	0	0	-
Science Advances	2,8	0	0	0	7	0	-
American Journal of Archaeology	2,6	0	0	8	0	0	A
The Anatomical Record	2,6	0	0	0	8	0	-

Journals	Score	WoS	Scopus	Dimen.	Open Alex	GS	ANVUR
Archaeological Prospection	2,4	9	0	0	0	0	A(2019)
Boreas	2,4	0	9	0	0	0	-
Heritage	2,4	0	0	0	0	9	A(2018)
The Journal of Hellenic Studies	2,4	0	0	9	0	0	A
Current Anthropology	2,2	0	10	0	0	0	A
Nature	2,2	0	0	0	10	0	A
African Archaeological Review	2,0	11	0	0	0	0	A
Journal of Computer Applications in Archaeology	2,0	0	11	0	0	0	-
Journal of Human Evolution	2,0	0	0	0	11	0	A
Advances in Archaeological Practice	1,8	12	0	0	0	0	-
Forensic Science International	1,8	0	0	0	12	0	-
Environmental Archaeology	1,6	0	20	0	0	18	A
Evolutionary Anthropology Issues News and Reviews	1,6	0	0	0	13	0	-
The South African Archaeological Bulletin	1,6	0	0	13	0	0	-
Dumbarton Oaks Papers	1,4	0	0	14	0	0	A
Ethnoarchaeology	1,2	15	0	0	0	0	-
International Journal of Osteoarchaeology	1,2	0	0	15	0	0	A
Journal of Agrarian Change	1,2	0	15	0	0	0	-
American Journal of Biological Anthropology	1,0	0	0	0	16	0	-
Archaeological Research in Asia	1,0	0	0	0	0	16	-
Studies in Conservation	1,0	0	0	16	0	0	A(2018)
Journal of Island and Coastal Archaeology	0,8	0	17	0	0	0	-
Journal of Mediterranean Archaeology	0,8	17	0	0	0	0	A
PLOS One	0,8	0	0	0	17	0	A
The Journal of Egyptian Archaeology	0,8	0	0	17	0	0	A

Journals	**Score**	**WoS**	**Scopus**	**Dimen.**	**Open Alex**	**GS**	**ANVUR**
Journal of the Economic and Social History of the Orient	0,4	0	0	19	0	0	-
Paleoamerica	0,2	20	0	0	0	0	-