www.ijarp.com ISSN 2456-9992 Page: 01-34

International Journal Advanced Research Publications

Research Article

Volume: 02
Issue: 07

BENCHMARKING SCHOLARLY IMPACT A COMPARATIVE BIBLIOMETRIC ANALYSIS OF RESEARCH PRODUCTIVITY ACROSS SCOPUS, WEB OF SCIENCE, ORCID, AND DIMENSIONS

[1]Amandeep Kaur, [2]Mamta, [3]Riju Kumari, *[4]Sunil Kumar

[1]Phd Research Scholar, Department of Library and Information Science Lamrin Tech Skill University Ropar, Punjab, India.

[2]Phd Research Scholar, Department of Library and Information Science HRIT University, Ghaziabad, UP, India.

[3]PhD Research Scholar Department of Library and information Sciences HRIT University Ghaziabad U P India.

[4]Chief Librarian, PIMS Medical College and Hospital Jalandhar Punjab India.

Article Received: 13 June 2026, Article Revised: 03 July 2026, Published on: 23 July 2026

***Corresponding Author: Sunil Kumar**

Phd Research Scholar, Department of Library and Information Science Lamrin Tech Skill University Ropar, Punjab, India.

Doi: https://doi-doi.org/101555/ijarp.2916

ABSTRACT

Background: Publication and citation indicators depend heavily on the databases from which they are derived; however, Scopus, Web of Science, Dimensions, and ORCID serve different infrastructural roles. This study aims to evaluate the strengths and weaknesses of the aforementioned platforms and aims to provide a design for researcher level comparison for the years 2020-2025. For this, a secondary evidence synthesis was conducted. Evidence for this synthesis was obtained from peer reviewed studies and from comparative studies and were documented mainly in the years 2020 to 2025. Evidence of a quantitative nature was described in the source studies. A primary data collection design was then described for study populations of between 100 to 300 researchers and 1,000 to 5,000 research outputs. The techniques and tools that were to be used for this design included DOI-first matching, normalized-title matching, researcher identifiers, metadata-completeness indices, non-parametric repeated-measures tests, agreement statistics, regressions, and sensitivity analysis. Evidence shows that multiple databases provide a substantial amount of coverage that is non-equivalent. One large multijournal study showed that 99.11% of the journals in the Web of Science were also in Scopus, in Dimensions were also in Scopus were also in Web of

Science, and that Dimensions had about 48.17% more journals than Scopus. ORCID facilitates strong identity linking and discoverability, yet profile completion and user adoption remain low. No single platform provides a complete or functionally equivalent representation of scholarly productivity. Multi-database triangulation, matching, field normalization, identity verification, and expert review are required for rigorous evaluation.Practical implications are as follows: This manuscript offers an auditable framework addressing the needs of librarians, universities, research offices, and evaluators. Originality/value: This manuscript creates a space for verified secondary benchmarks to be accessed, thus eliminating the need for fabricated metrics that rely on proprietary exports while allowing for a succinct and comprehensive research design.

KEYWORDS: Bibliometrics; Scientometrics; Scopus; Web of Science; Orcid; Dimensions; Research Productivity; Citation Analysis.

1. INTRODUCTION

The evaluation of research now functions through an extensive network of bibliographic databases, persistent identifiers, institutional research information systems and analytic dashboards. Measures such as publication counts, citation counts, h-indices, indicators of collaboration, and field-normalized metrics are often considered to be fixed attributes of a researcher or an institution. In reality, they are merely snapshots produced by a specific database, at a specific moment, and under a particular policy of indexing. Changing the source may result in a new set of documents, a new citation graph, a new classification of fields, a new treatment of document types, and a new attribution of document authors. Hence, a database is not a neutral container, but an integral part of a measurement system (Visser et al., 2021; Stahlschmidt& Stephen, 2022).

As commercial, multidisciplinary citation databases, Web of Science and Scopus the most established systems. Both curate lists of sources, design citation networks, and provide author and institutional evaluative analytic profiles. Their coverage is extensive, but, unaligned. Differences impact citation of conference proceedings, books, regional journals, and non-English publications, as well as works in fields whose communication practices do not adhere to a journal-centric system. Recent analyses also reveal that citation indicators, which may appear similar, can diverge as each database crawls a distinct citation network and applies disparate data processing policies (Pranckutė, 2021; Singh et al., 2021).

Dimensions extended the analysis by associating publications with grants, patents, clinical trials, policy documents, datasets, and researchers. Its broader ingestion philosophy may enhance recall and facilitate cross-domain analysis. However, describing platforms with broader coverage as better is misleading. A platform that possesses more records may also have more irrelevant records, more heterogeneous document types, and varied reference completeness. On the other hand, a platform that is more selective may be more homogeneous but may also contain all relevant, local, or applied research. The relevant question is therefore not which database is the most satisfactory database overall, but which database or combination of databases is the most satisfactory for the particular purpose of the research assessment.

ORCID involves a different analysis altogether. This is not a citation-indexing database. ORCID is a persistent person-identifier with an integrated system that connects researchers to funding, positions, peer reviews and other research outputs. Its merit consists in the resolution of research identities, and interoperability and control by the researcher. Considering ORCID citation coverage like that of Scopus, Web of Science, or Dimensions, is misleading. ORCID permits better matching and decreases name ambiguity, but an ORCID record is only as complete as its adoptions, permissions, integrations, and researcher involvement (Teixeira da Silva, 2020; Wang et al., 2024; Porter et al., 2025).

These functional discrepancies bring about real consequences. Indicators based on Scopus or Web of Science may be required for promotion, accreditation, performance-related funding, and ranking-related submissions. Indicators specific to Scopus or Web of Science may be used by researchers to capture and present their impact. Integration of ORCID by research and library services may be due to its inclusion in a CRIS. When custom indicators are not reconciled, the same researcher may have different totals for publications, citations, h-indices, collaboration profiles, and open access rates in different systems. When metrics are used as a threshold instead of providing context, the discrepancies may result in serious issues.

This paper addresses two needs. First, it provides a summary of comparable quantitative findings based on published literature from 2020 to 2025 that can be cited. Second, a direct comparison of researchers at this level using matched data from Scopus, Web of Science, Dimensions and ORCID is facilitated by this paper in a way that can be readily analyzed. This paper intentionally does not report unpublished record-level data due to the absence of proprietary data. The secondary synthesis provides only data from verified sources, and all proposed inferential analyses are placed in the realm of the primary data. This separation is crucial for the reproducibility and rigor of the methods described.

This paper captures the documented discrepancies in coverage of research outputs by the platforms, features of the platforms, visibility of citations and completeness of metadata, researcher identification, and researcher record deduplication, and provides a comprehensive bibliometric and statistical framework. This paper captures the consequences of choosing a specific research evaluation database, and proposes a prudent model that incorporates multiple research evaluation databases. The primary research questions relate to the coverage of research outputs by the platforms, the visibility of citations, the completeness of metadata, and the extent to which platforms endorse ORCID, and the discipline(s) of the research, as well as the consequences of using a single platform.

1.1 Research Aims

This research aims to do the following:

1. Analyze the coverage of publications and journals by Scopus, Web of Science and Dimensions, and compare findings for the 2020–2025 evidence interval.
2. Analyze the differences that may be reported on citation-link coverage and also normalized citation impact across the different citation-indexing services.
3. Analyze the coverage, and also the degree of coverage, of metadata bibliographic data that are used for matching across different databases.
4. Analyze ORCID, not as a citation database, but as an identity, attribution, interoperability, and visibility infrastructure.
5. Formulate a reproducible matching protocol for the benchmarking of researchers and research institutes.
6. Propose a multi-source, Responsible Research Assessment framework.

1.2 Research Queries

This research seeks to answer the following:

7. What are the differences in coverage and overlap across documented coverage of Scopus, Web of Science, and Dimensions?
8. What are the citation-link and impact differences that may be reported across the citation-indexing services?
9. What metadata and identifiers are critical for matching across research platforms?
10. How reliable are ORCID records for author-level assessments?

11. What is the appropriate statistical modeling for bibliometric data that has a matching, skewed, and zero-inflated distribution?
12. What suite of tools and services best enables a fair and reproducible assessment?

1.3 Primary Matched-Export Study Predictions

The following predictions, which cannot be validated retrospectively against the plethora of studies that have been published, are put forth for when matched record exports becomes available:

H1: There will be significant differences in coverage of publications by Scopus, Web of Science, Dimensions, and outputs linked to ORCID.

H2: There will be significant differences in citation counts for Scopus, Web of Science, and Dimensions for published works that are matched.

H3: There will be significant differences in the completeness of metadata across different platforms.

- H4: Researcher-level publication and citation statistics are positively correlated across all platforms but show a lack of perfect correlation.
- H5: Discipline, document type, language, presence of a DOI, and open-access status are all predictors of multi-platform indexing.
- H6: The presence of a DOI positively impacts cross-database matching.
- H7: The completeness of an ORCID profile positively impacts the number of verified outputs linked to the profile.

2. Literature Review

Since 2020, large-scale comparisons have focused less on simple size comparisons of the databases and more on document overlap, citation-link completeness, and fitness for specific assessment tasks. Visser et al. (2021) conducted a broad-scale comparison of Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic, and demonstrated that the differences in coverage were dependent on temporal, disciplinary, and document type selectors. The main innovation in their methods was that citation-linked document matching and citations must be the primary focus of database comparisons. A broad count of databases may not capture whether matched records are scholarly, contain citations, and are fit for the intended task.

Singh et al. (2021) presented a comprehensive quantitative analysis of journal coverage across Web of Science, Scopus, and Dimensions. Their work demonstrated that the overlap among databases was very high, with substantial overlap among Web of Science and Scopus. However, Dimensions has considerably more journals. Significant overlap coupled with significant unique coverage is a case of high synergy along the core with meaningful difference of coverage along the margins. Such margins are likely to be impacted by language, geography, discipline, and source selection.

Citation coverage demonstrates a similar core-periphery pattern. In the study by Martín-Martín et al. (2021), over three million citations across 2,515 highly-cited documents were examined. Dimensions captures most of the citation links that Scopus and Web of Science contain, but coverage varies by discipline and is consistently poor in many areas of the humanities. Thus, citation counts are not synonymous. An author's research impact is not objectively assessed based on the quality of their work, which can be evidenced by the difference in citing documents.

Guerrero-Bote et al. (2021) analyzed Scopus and Dimensions at the country and institutional levels and illustrated that large datasets do not entail the same representation at the country or institutional levels. This point was elaborated on by Stahlschmidt and Stephen (2022), who analyzed the structures of citation networks and normalized impacts in Web of Science, Scopus, and Dimensions, and showed that for the 41,848 publications that were common to the three databases, there were system-specific variations in normalized impact, especially in relation to applied research. The outcome of this study indicates that field normalization fails when there are system differences in the citation networks and the reference networks, and citation networks.

Pranckutė (2021) compiled the long-standing Web of Science-Scopus discussion and argued that coverage, ease of use, metrics, ease of export, and the research question all determine what database to use. The review noted that the coverage of books and conference proceedings is especially problematic for fields that rely heavily on these outputs, a source of major inequity across fields. Evaluation frameworks that rely on journal-based databases neglect the arts, humanities, computer science, and social research, and locally focused social research, even when the databases are operating as intended.

Separating metadata quality from breadth means considering other dimensions. For cross-platform matching, accuracy, normalization, author order, journal codes, volume, pagination, affiliations, and reference lists are critical. biblioverlap, an R package developed by Vieira and Leta (2024), responds to a need for matching bibliographic datasets, with a focus on the auditable cross-source integration. Culbert and colleagues (2025) considered 16.8 million publications from OpenAlex, Web of Science, and Scopus. They reported that, although reference coverage was similar across the three databases, OpenAlex had more coverage of authors' ORCID identifiers, provided an average of fewer authors' abstracts, and had a coverage of open access indicators that was similar to that of the other databases. The main point is that no single platform is the best for addressing every metadata issue.

The analysis of funding metadata related to the same publication can also provide different results. Kokol and colleagues (2023) analyzed the coverage of funding information in Scopus and in Web of Science, and in the coverage of the selected medical journals, platform-dependent discrepancies were observed. Funding acknowledgements provide evidence of the sponsors of research and can indicate the outcomes of research, and incomplete coverage can distort the analysis of sponsors. The same problems can be identified with authors' affiliations and with the normalization of organizations, especially when institutions merge, change names, or publish in different languages.

ORCID is a solution for the identity layer, not the citation layer. While acknowledging ORCID's merits, Teixeira da Silva (2020) expressed concerns about the involuntary use of ORCID, data quality, and research integrity. Wang et al. (2024) explored the potential of using ORCID for curriculum vitae (CV) analysis and identified inconsistencies in the extent to which ORCID profile fields are populated. Porter et al. (2025) explored the use of ORCID, showing that individual and contextual factors influence the extent of ORCID registration. Houghton and Foster (2024) explored the perception of ORCID among academics and identified resistance to the use of ORCID, perceiving it as an administrative task and a form of surveillance. Collectively, these studies demonstrate that it is a misconception to think that a publication is absent when it is not linked to ORCID.

Notably, use of ORCID improves the ability to discover and attribute research. Ibrahim and Ramzy (2025) studied the use of ORCID among 569,038 publications from Arab universities in the Web of Science and employed propensity-score matching with 17 covariates. They identified positive associations between the use of ORCID and publication, use, and citation, and were the first to provide evidence of the benefits of using ORCID for discovering and citing research.

Author-name disambiguation continues to be a major issue in bibliometrics. Names may be changed, initials may be used, names may be altered through transliteration, names may be compound, and affiliations may be fluid. For example, researchers may be split into different profiles or merged into one due to the nature of their names and fluid affiliations. One example of this problematic nature is seen in the work of Kim et al. (2021) as they utilize ORCID to create labeled data sets to test author disambiguation methods. Also contributing to this area of research is Cappelli et al. (2025). They focus on deep learning and cite the need for persistent identifiers and improved metadata. They state that persistent identifiers alone cannot achieve reliable author-level benchmarks. Such benchmarks must also incorporate DOIs, authorship, and topical data, and be supported by a manual review for ambiguous cases.

Greater attention has been paid to outlining methods. This is exemplified in the bibliometric research of Donthu et al. (2021), Mukherjee et al. (2022), Moral-Muñoz et al. (2020), and Öztürk et al. (2024). Research by Lim et al. (2024) advocates for the combination and cleansing of data in advance of the application of multiple tools. Montazeri et al. (2023) attempt to establish the minimal requirements for reporting in biomedicine in the context of bibliometrics, while Page et al. (2021) and Rethlefsen et al. (2021) advocate for tenets of transparent search and selection. Together these published works advise that transparent data retrieval methods and thorough data cleansing are of utmost importance and cannot be substituted for with aesthetically pleasing data visualization.

Responsible assessment frameworks include normative constraints. The Leiden Manifesto proposes that quantitative evaluations should be used to complement expert evaluations, not to replace them (Hicks et al., 2015). DORA advocates against the use of bibliometric journal rankings to assess the quality of single articles or researchers. Moed (2020) and Thelwall (2020) argue that bibliometric indicators should only be used with respect to the mission of the institution, the academic field, the stage of the career, and the data constraints. This study supports Moed (2020) and Thelwall (2020) by assuming that a bibliographic database comparison is the basis for fair interpretation of the results.

Several reviews, including Aria and Cuccurullo (2017), have proposed bibliometrics workflows, while Passas (2024), González-Alcaide (2021), and Hulland (2024) noted that bibliometric reviews should involve manual design steps and cannot rely on automated charting. Because citations are not evenly distributed, a journal-based median is preferable to an average (Kiesslich et al., 2021). Narrative bibliometrics and responsible assessment push for contextual interpretation that is transparent and sophisticated, rather than simplified ranking (Coalition for Advancing Research Assessment, 2022; Torres-Salinas et al., 2024; UNESCO, 2021).

The primary research gap is the tendency to treat four disparate systems as one in a league table. When direct comparisons are made, they are only made on two or three citation databases. In many cases, ORCID is either left out or included as just another profile source. The design must separate indexing and identity, and determine which of the hypotheses are dependent on matched gated data. The current paper aims to bridge this gap by using verified comparative data and a comprehensive protocol that prescribes detailed steps following the provision of data exports.

3. Conceptual Framework

A framework that conceptualizes scholarly productivity incorporates five interrelated layers. The first layer is platform design, which includes proprietary ownership, source selection, document design, updates, citation construction, and access; the second layer is the observable quality of data, which includes coverage, metadata, persistent identifiers, author affiliations, document types, and reference links; the third layer is the metrics, which includes publication counts, citation counts, h-indices, collaboration metrics, and normalized impact; the fourth layer is reconciliation, which includes removal of duplicates, matching, field normalization, and a challenge to the data's validity; and the fifth layer, the use of the data, combines quantitative data with qualitative assessments, and is framed by the institution's mission, the discipline, and a theory of responsible assessment. When there is a distortion or bias at the beginning of the process, all subsequent measures will be affected. Thus, all layers must be documented.

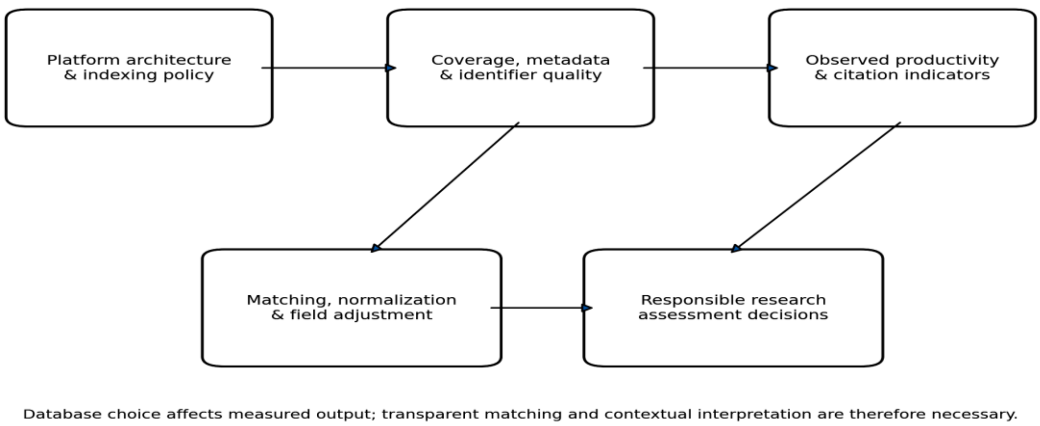

Figure 1. Conceptual framework linking platform architecture to evaluation outcomes. Source: Author-developed synthesis.

4. MATERIALS AND METHODS

4.1 Research Design

The completed manuscript employs a two-part design. A synthesizes secondary evidence, mainly focused on the years 2020-2025, from peer-reviewed articles and comparative studies, as well as the documentation from authoritative platforms. B presents a prospectively defined protocol for direct matched-record analysis of researchers and their outputs for the period of 1 January 2020 to 31 December 2025. The separation was done to avoid the confusion of the original platform-export results with published benchmarks.

4.2 Evidence Identification and Eligibility

Through specific searches evidence was identified on the comparative coverage, citation, and metadata; author disambiguation; the adoption of ORCID; bibliometrics; and responsible research assessment. The main focus was on peer-reviewed articles that contained datasets, clear methods, and quantitative results, as well as the documentation that described the functions of the platforms. Research was included if it analyzed at least one of the platforms of interest and addressed coverage, citation, metadata, identity, and matching or evaluation. Methodological promotions that lacked transparency were considered inappropriate and were removed from the corpus. Due to the lack of homogeneity of outcomes, units, time periods, and denominators, this corpus was a structured base of evidence and not a formal meta-analysis.

Table 1. Functional distinction among the four platforms.

Platform	Primary role	Core analytical value	Key caution
Web of Science Core Collection	Curated bibliographic and citation index	Selective source coverage, cited-reference searching, author/institution analytics	Metrics reflect its curated source and citation universe
Scopus	Curated abstract and citation database	Broad multidisciplinary coverage, profiles, citations, affiliation and journal metrics	Coverage and citation totals are not identical to Web of Science
Dimensions	Linked research-information and citation platform	Broad publications plus grants, patents, trials, policy and related entities	Broad ingestion requires careful filtering and matching
ORCID	Persistent researcher identifier and profile infrastructure	Identity resolution, attribution, portability and interoperability	Not a conventional citation database; profile completeness is user/integration dependent

Note. Platform descriptions reflect functional roles documented by the platform providers and evaluated in the comparative literature. ORCID is intentionally not ranked as a citation index.

4.3 Proposed Unit of Analysis and Sampling

The ideal primary unit is the researcher (nested in department and institution) with their publications tracked as repeated observations. An adequate sample is believed to comprise 100–300 active researchers from all the sciences and fields. To achieve a stratified sample, the sample ought to reflect balance by discipline, career stage, and institution. Strong identifiers or a verified combination of bibliographic fields will enable identification of publications in the time frame 2020–2025. The anticipated analytical corpus is 1,000–5,000 unique outputs in a given framework, following cross-platform deduplication.

4.4 Search and Export Strategy

To enumerate researchers, the most reliable source of institutional lists should be consulted. For each researcher, the following should be recorded (in order): (a) ORCID iD, (b) Scopus Author ID, (c) ResearcherID (Web of Science), (d) Researcher ID (Dimensions), (e) name variations, and (f) history of affiliations. Search the platforms in the order given above for each identifier. All other searches may be conducted as name–affiliation searches. Each export should document the exact search criteria, retained filters, export timestamp, and the cited fields. Because citation counts change over time, platforms must be searched in a time optimal manner. The date each search was conducted should be recorded at the citation level.

4.5 Record Matching and Deduplication

Matching is performed in steps. The most powerful publication-level key is the exact normalized DOI. Normalization of the DOI involves the removal of protocol prefixes, resolver domain, whitespace, trailing punctuation, and case sensitivity. When a record lacks a DOI, it is matched using the normalized citation via the publication year, the first author, the source's title, volume, issue, pages or article number, and the co-author's pattern. When attempting to match titles, exact match is preferred, otherwise, fuzzy matching of the title with token-set similarity is used. A threshold of 95 is considered an automatic acceptance, 90 - 94 is accepted for a manual review, whereas a score below 90 is considered a non-match in absence of validation from other fields. Variants of conference papers and early access articles undergo a separate review due to differences in titles and publication years. A confidence code and an audit note are provided for each matched record.

4.6 Metadata Completeness Indices

The computation of the completeness of metadata for bibliographic sources is done using the equation MCI = 100 * (fields present)/(fields assessed) and involves the assessment of data on the following: DOI, title, full author list, full author affiliations, abstract, author keywords, source, document, and metadata on the publication, citation and reference, OA status, funding, researcher ID, and subject. In the primary analysis, all fields are assigned equal weight. In the secondary analysis, fields may be weighted differently. ORCID is assessed through a separate Profile Completeness Index and includes a verified domain email, or other trusted affiliations, along with an assessment of employment, education, associated peer reviewed and other works, and funding. Completeness of data is assessed for each individual field and is not hidden in an aggregated score.

4.7 Bibliometric Indicators

Researcher-level indicators include total number of publications, total number of citations, median citations, mean citations, interquartile range, h-index, i10-index, share of research outputs that are uncited, annual growth of research outputs, the percentage of research outputs that are open-access, the rate of international collaboration, the number of sources that indexed the researcher's works (one, two, three, or four sources), among other indicators. Document-level indicators include the presence of the document in four platforms, total citations by source, presence of DOI, open-access status, type of collaboration, funding acknowledgment, field of research, and a completeness score for the metadata. Since citation distributions are right-skewed, means should be complemented by medians and percentiles as well as log-transformed or count-regression models.

4.8 Statistical Analysis

Descriptive analysis must include a full account of skewness and kurtosis as well as 95% confidence intervals. Shapiro–Wilk tests should be used to assess publication status across four platforms, which is a repeated binary outcome, along with the histograms and Q–Q plots. Frequencies may be analyzed using Cochran Q test and subsequently by McNemar test with Holm correction. Citation counts for Scopus, Web of Science, and Dimensions should be analyzed using Friedman test and paired Wilcoxon signed-rank test, as citation counts are skewed and contain ties. Spearman rank correlation should be employed for citation counts across all platforms. Intraclass correlation should be used along with a Bland–Altman plot to assess the degree of agreement across the citation platforms.

To examine the predictors of multi-platform coverage, mixed-effects logistic regression can be performed, having publications nested within researchers or institutions. Due to the expected overdispersion, citation counts can be modeled using negative-binomial regression.

If there are a substantial number of outputs that have never been cited, a zero-inflated model can be used. Some possible predictors to be included are the presence of a DOI, publication accessibility, the year of publication, type of document, level of collaboration (international), funding, publication language, field of research, and the journal. Regarding the presence of metadata, Friedman or mixed-effects ordinal/linear regression can be used to assess differences, depending on the distribution of the scores. Model diagnostics and multiple-testing corrections should always be provided, along with effect sizes, confidence intervals, and exact p-values.

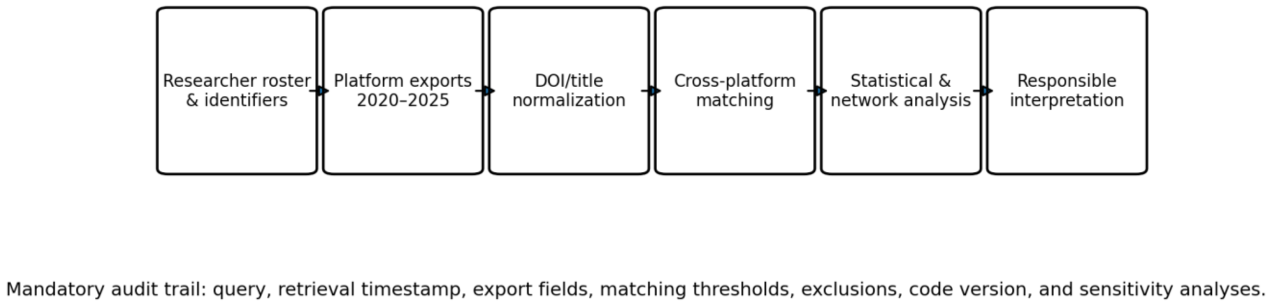

Figure 2. Reproducible primary-data pipeline. Source: Author-developed protocol.

Table 2. Verified evidence corpus used in the secondary synthesis.

Source	Platforms	Unit/method	Contribution
Visser et al. (2021)	Scopus, WoS, Dimensions, Crossref, Microsoft Academic	Large-scale document and citation-link coverage	Coverage varies by field, year, and document type.
Singh et al. (2021)	WoS, Scopus, Dimensions	Journal lists and country outputs	High shared core; Dimensions broader; WoS more selective.
Martín-Martín et al. (2021)	Scopus, WoS, Dimensions and others	3,073,351 citations to 2,515 documents	Citation coverage differs across 252 categories.
Guerrero-Bote et al. (2021)	Scopus, Dimensions	Country- and institution-level outputs	Global breadth does not produce identical local representation.
Pranckutė (2021)	WoS, Scopus	Critical comparative	Database choice depends

Source	Platforms	Unit/method	Contribution
		review	on purpose, coverage, metrics, and export limits.
Stahlschmidt& Stephen (2022)	WoS, Scopus, Dimensions	Citation networks; 41,848 shared publications	Normalized impact remains database dependent.
Kokol et al. (2023)	WoS, Scopus	Funding metadata in medical journals	Funding-information coverage differs by platform.
Wang et al. (2024)	ORCID	Academic-profile data availability	Profile fields are unevenly complete.
Vieira & Leta (2024)	Multiple bibliographic datasets	Document matching and R package	Matching must be explicit, normalized, and auditable.
Houghton & Foster (2024)	ORCID	Academic and librarian perceptions	Adoption may be constrained by burden and monitoring concerns.
Culbert et al. (2025)	OpenAlex, WoS, Scopus	16.8 million shared publications	Reference coverage comparable; metadata strengths differ.
Porter et al. (2025)	ORCID	Researcher adoption	Adoption is patterned and should not be assumed universal.
Ibrahim &Ramzy (2025)	ORCID in WoS metadata	569,038 publications; propensity matching	ORCID availability associated with greater usage and citations.
Cappelli et al. (2025)	Author-disambiguation literature	Systematic review	Persistent identifiers complement algorithmic disambiguation.
Nowakowska (2025)	Bibliometric datasets	Preprocessing framework	Cleaning decisions materially affect analysis.
Scheidsteger et al. (2025)	WoS, Scopus, Dimensions, OpenAlex	Field-normalized impact comparison	Normalized scores remain source sensitive.

Table 3. Quantitative benchmarks extracted from comparative studies.

Source	Benchmark	Reported value	Interpretive boundary
Singh et al. (2021)	WoS journals also in Scopus	99.11%	Journal-level master-list comparison
Singh et al. (2021)	WoS journals also in	96.61%	Journal-level master-list

Source	Benchmark	Reported value	Interpretive boundary
	Dimensions		comparison
Singh et al. (2021)	Scopus journals also in Dimensions	96.42%	Journal-level master-list comparison
Singh et al. (2021)	Dimensions additional journals vs WoS	82.22%	Relative breadth reported by authors
Singh et al. (2021)	Dimensions additional journals vs Scopus	48.17%	Relative breadth reported by authors
Martín-Martín et al. (2021)	Scopus citations captured by Dimensions	84%	Highly cited multidisciplinary corpus
Martín-Martín et al. (2021)	WoS citations captured by Dimensions	88%	Highly cited multidisciplinary corpus
Stahlschmidt& Stephen (2022)	Publications shared by all three databases	41,848	Normalized impact comparison
Culbert et al. (2025)	Recent publications shared by three databases	16.8 million	Cleaned reference-coverage comparison
Ibrahim &Ramzy (2025)	Arab-university publication dataset	569,038	WoS records, 2019–2023
Ibrahim &Ramzy (2025)	Covariates used in propensity matching	17	Adjustment for selection and confounding

Table 4. Variables for the proposed matched-record dataset.

Domain	Variables	Level	Coding/derivation
Identity	Researcher ID, ORCID, Scopus ID, ResearcherID, Dimensions ID, name variants	Researcher	Validated identifiers and affiliation history
Bibliographic	DOI, title, authors, source, year, volume, issue, pages/article number	Publication	Raw plus normalized values
Coverage	Present in Scopus/WoS/Dimensions/ORCID	Publication	Binary 1/0 with retrieval date
Impact	Citation counts, h-index, i10, uncited status	Publication/researcher	Platform-specific; never merged without label
Metadata	Abstract, keywords, affiliations, funding, references, OA, identifiers	Publication	Field presence and completeness score

Domain	Variables	Level	Coding/derivation
Context	Field, institution, country, language, document type, collaboration	Publication/researcher	Harmonized controlled categories
Audit	Query, export file, timestamp, match method, confidence, reviewer	Record	Required for reproducibility

Table 5. Statistical analysis plan for primary platform exports.

Question/outcome	Recommended method	Effect/agreement statistic	Key assumption or diagnostic
Platform presence across four sources	Cochran Q; pairwise McNemar	Odds ratios; matched proportions	Same publications/researchers measured repeatedly
Citation counts across three sources	Friedman; paired Wilcoxon	Kendall W; rank-biserial effect	Skewed paired counts; ties handled
Researcher indicator association	Spearman correlation	rho with 95% CI	Monotonic relationship
Absolute agreement	ICC and Bland–Altman	ICC; mean bias; limits of agreement	Inspect heteroscedasticity and scale
Metadata completeness	Friedman or mixed-effects model	Kendall W / standardized beta	Score distribution and repeated records
Multi-platform indexing	Mixed-effects logistic regression	Adjusted odds ratios	Nested publications and researcher random effects
Citation visibility	Negative-binomial regression	Incidence-rate ratios	Overdispersion; zero inflation
Sensitivity	DOI-only, article-only, outlier-trimmed, field-normalized	Change in estimates	Pre-specified alternative definitions

Aiding performance analysis will be Bibliometrix/Biblioshiny within R, and for network visualization, either VOSviewer or Gephi. Python or R will be necessary for data cleaning and statistical modeling, along with OpenRefine for authority normalization. Auditable inspection in Excel will be the only use of the program, and will refrain from using it as the main engine for analysis. Methodology will dictate software and tools selection will be in accordance with an analysis scheme. Package versioning along with session information will be preserved. The benchmarks for bibliometric reporting as well as recent forms of bibliometric reporting critique (Montazeri et al., 2023; Koo& Lin, 2023; Li et al., 2024) will be the basis for reporting.

4.9 Ethical and Reproducibility Considerations

Bibliographic metadata along with researcher information that is public, or licensed to the institution, will comprise the data of the study. In accordance with ethical data practice, the data will be secured and unidentifiable researchers will not be ranked publicly. The visibility settings of ORCID will be honored. When licensing allows, analysis plans, along with the code, a variable dictionary, matching rules, and the results in their aggregate form will be made available. Otherwise, unpublished conclusions will be drawn.

5. RESULTS

5.1 Nature of the Results

The following results will be a secondary synthesis from verified, published, and publicly accessible research. These results will not be comparable to an unpublished database export. Any analysis requiring a more granular export of data for record-level concordance, such as Friedman tests of citation counts, co-authorship networks, citation analysis, h-index comparison, Bland-Altman analysis, or regression, will be restricted to a future date when source data have been published. This restriction is a methodological limitation, and is not to obfuscate missing data.

5.2 Coverage and Overlap

The most comparable data come from Singh et al. (2021). According to the study, 99.11% of journals in Web of Science were also in Scopus, and 96.61% were in Dimensions. Singh et al. also report that 96.42% of Scopus journals were in Dimensions. For mean inclusion (sample SD 1.50; range 96.42–99.11), a mean of 97.38% was calculated across the three. While the average suggests a significant degree of overlap, the mean should not be interpreted as the inclusion in the three, since the journals included in the three comparisons are not the same, and at the same time, unique sources can impact the field, language and country profiles.

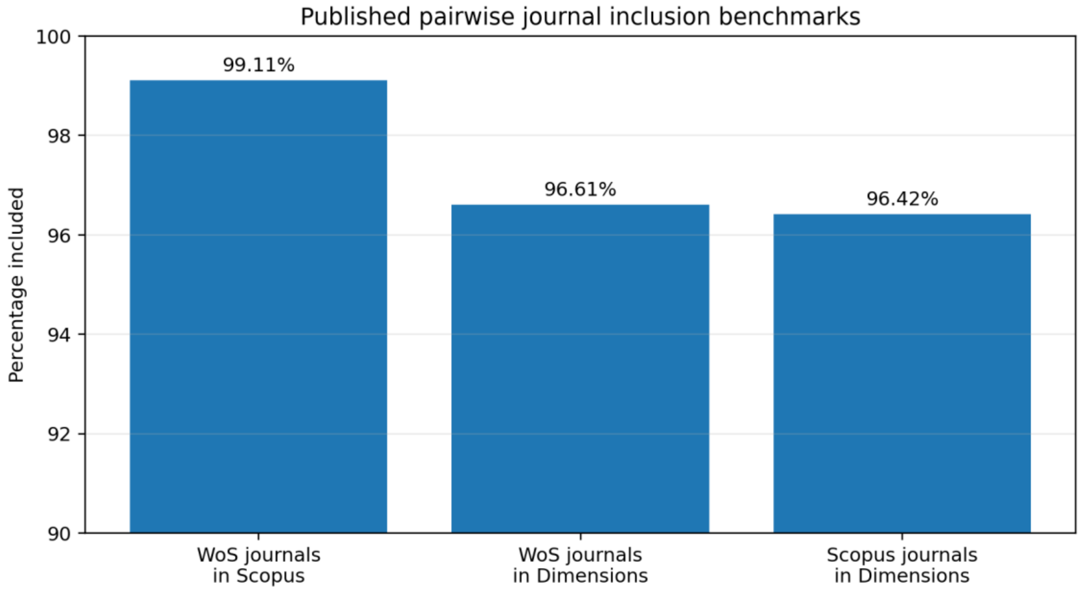

Figure 3. Published pairwise journal inclusion benchmarks. Data source: Singh et al. (2021).

Dimensions was reported to cover 82.22% more journals than Web of Science and 48.17% more than Scopus. The mean relative advantage across these two comparisons was 65.20% (sample SD 24.08). The wide difference between the two percentages reflects Scopus's broader source coverage relative to Web of Science while confirming the more expansive Dimensions universe. Breadth, however, is only one quality dimension; evaluators must still consider curation, document type, metadata, and relevance.

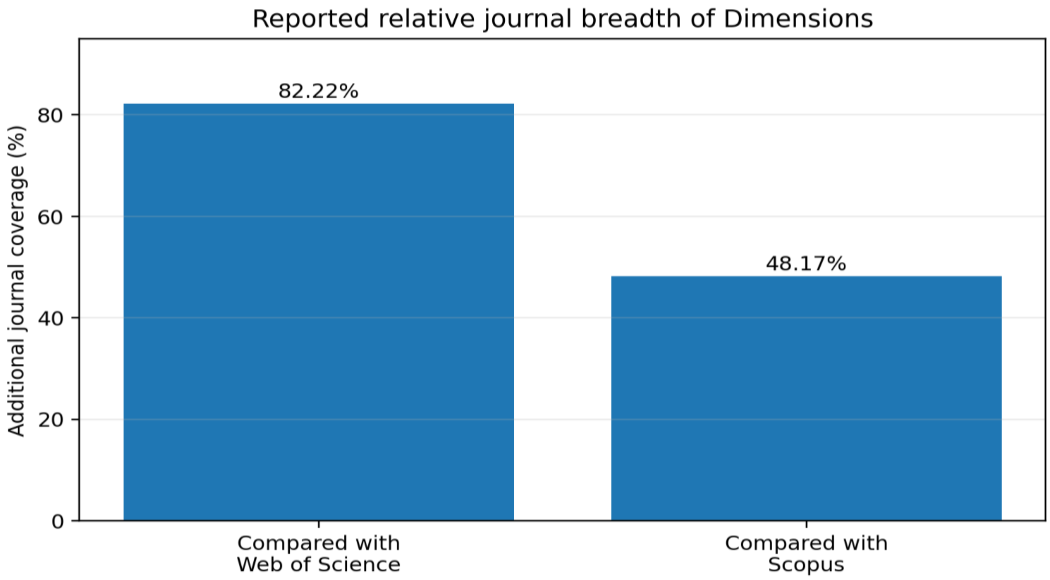

Figure 4. Reported relative journal breadth of Dimensions. Data source: Singh et al. (2021).

5.3 Citation-Link Coverage and Impact

Martín-Martín et al. (2021) analyzed 3,073,351 citations to 2,515 highly cited English-language documents across 252 subject categories. Dimensions contained 84% of the citations found by Scopus and 88% of those found by Web of Science, an average inclusion of 86.0% (sample SD 2.83). Dimensions found more citations than Web of Science in 185 categories but more than Scopus in only 36, illustrating that aggregate coverage and field-level performance can diverge. The reported humanities gaps reinforce the need for disciplinary stratification.

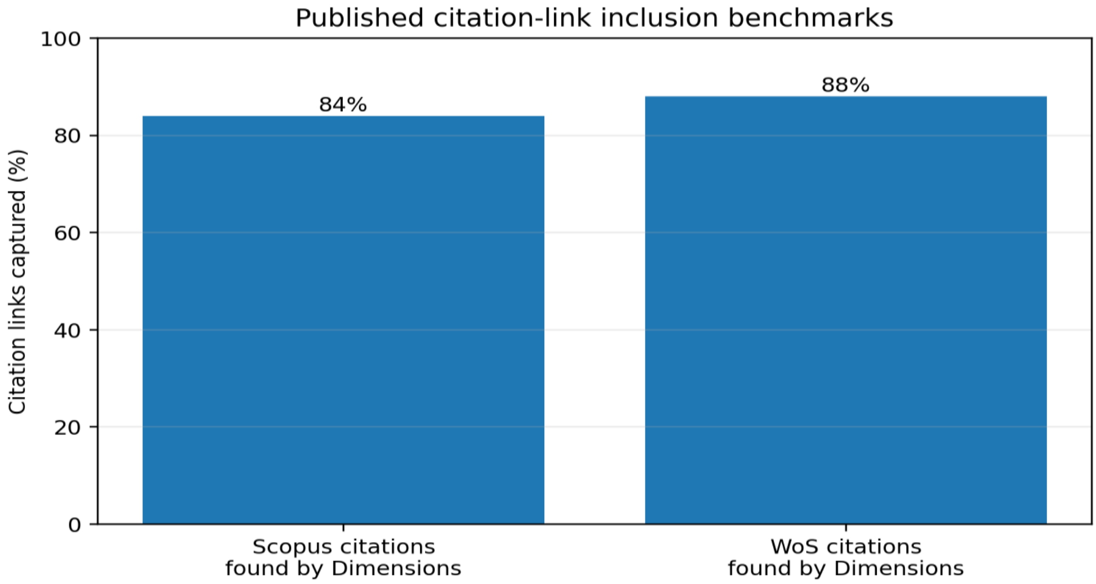

Figure 5. Published citation-link inclusion benchmarks. Data source: Martín-Martín et al. (2021).

Stahlschmidt and Stephen (2022) examined 41,848 publications from three commercial databases and determined citation impacts, which were still disparate. The applied-research sectors were stronger (in relative terms) in Scopus and Dimensions, as opposed to Web of Science. This conflicts with the claim that field citation normalization, in effect, abolishes database implications. Normalization concerns cite reference groups, but the reference group is database dependent.

5.4 Metadata Completeness and Matching

Studies dealing with the evaluation of metadata are complementary in nature, instead of singularly winning. Culbert et al. (2025) examined 16.8 million publications from OpenAlex, Web of Science, and Scopus. Their average reference counts, and internal reference coverage, were similar, but metadata differed and were field dependent. OpenAlex, for example, contained more ORCID iDs and had fewer abstracts, but had an equal number of open-access indicators. This finding is important, as it shows that a database may be strong in identifiers, and in openness, but weak in descriptive text. Adding to this, Vieira and Leta (2024) showed that precise normalization and matching software had to be employed in order to conduct a cross-database analysis. Any simplistic form of cross-database analysis, such as title matching, was highly discouraged.

According to the ORCID evidence, there are two findings to mention. First, discoverability and identity linkage improves with ORCID. In an empirical study by Ibrahim and Ramzy (2025), 569,038 Web of Science publications were used, and by employing a propensity-score matching technique which included 17 covariates, significant positive effects were found regarding the availability of ORCID on usage and citation ($p < .001$). Second, ORCID cannot be presumed to be complete. Wang et al. (2024) talked about the inconsistent availability on profile fields and Porter et al. (2025) talked about inconsistent adoption. Through their research, Houghton and Foster (2024) had noted the uneasiness regarding the burden of administration and the constant need to check. Because of this, the ORCID-linked output count should be seen best as a completeness measure of linkage and not as a total count of the researcher's publications.

5.6 Evidence Distribution and Functional Roles

The verified corpus for this manuscript was coded by evidence domain. Eight sources provided evidence on methods or responsible assessment. Seven sources provided evidence on coverage and overlap; six sources provided evidence on metadata and matching; six sources provided evidence on ORCID or identity and five sources provided evidence on citation impact. Due to the nature of the field, overlaps in the categories were expected. The strongest areas of literature were shown to be coverage and methods; however, there are still a lack of direct researcher-level studies in the four-platform literature.

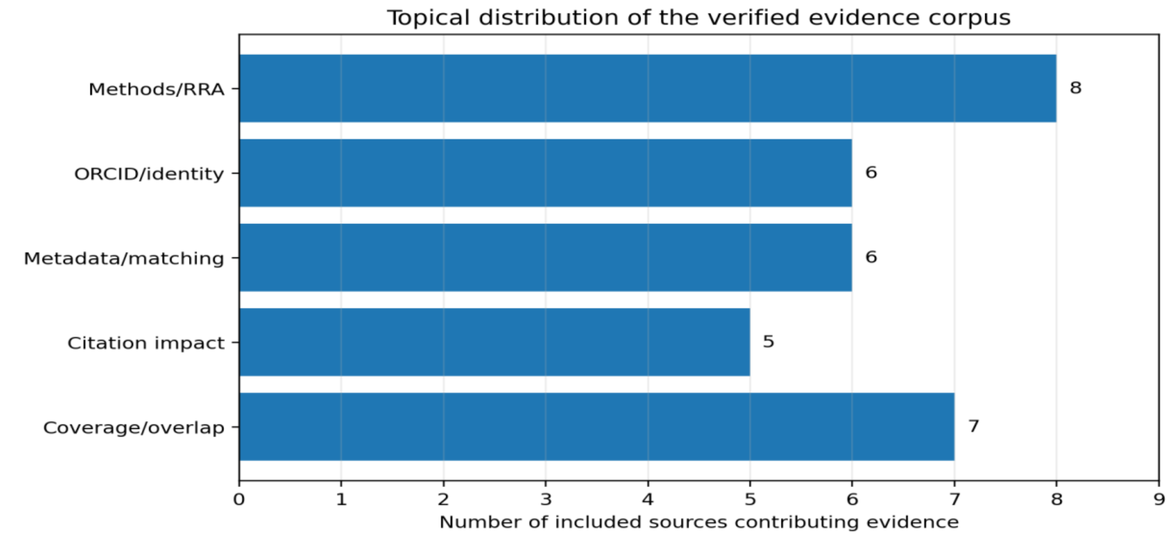

Figure 6. Topical distribution of the verified evidence corpus. Categories are non-exclusive.

An evidence-informed functional matrix was created to clarify role differentiation. The scores are ordinal interpretive ratings, not platform metrics. Web of Science and Scopus receive the strongest role in citation analytics; Dimensions receives the strongest role in breadth and linked metadata; ORCID receives the strongest role in identity and interoperability. The zero assigned to ORCID citation analytics indicates non-applicability, not poor performance.

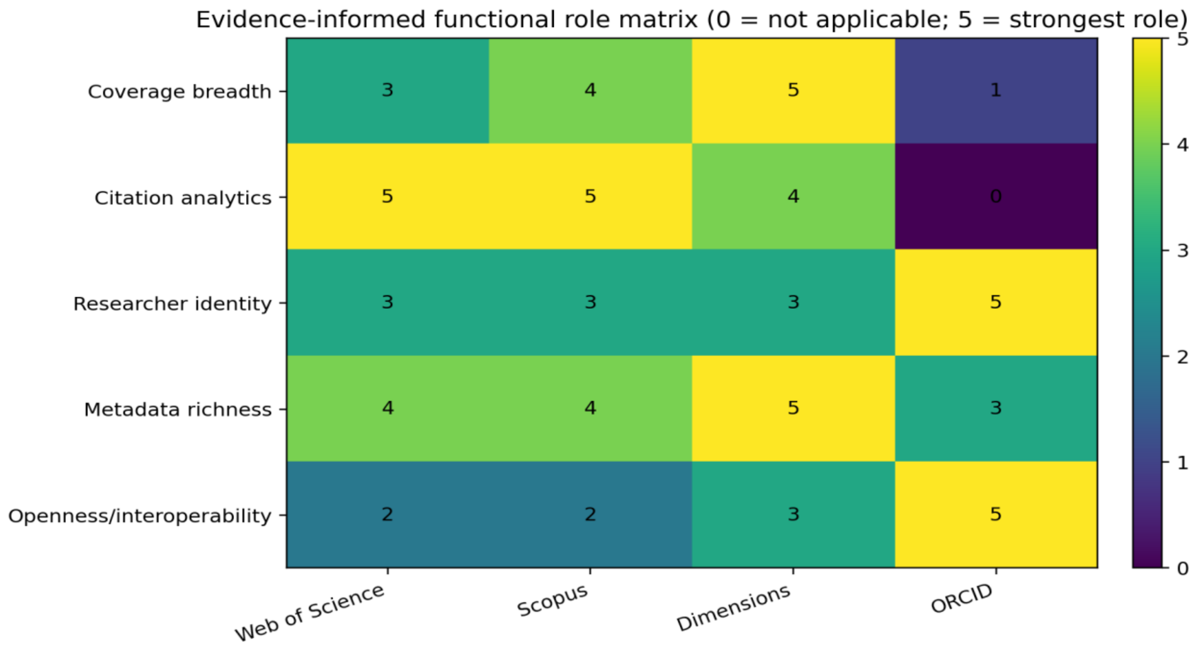

Figure 7. Evidence-informed functional role matrix. Scores summarize role strength and are not empirical platform-quality measures.

Table 6. Evidence-synthesis findings by evaluation domain.

Domain	Synthesis finding	Confidence	Implication
Journal/publication breadth	Large shared core coexists with meaningful unique coverage; Dimensions is generally broader and WoS more selective.	High	Report overlap and unique records, not only totals.
Citation visibility	Citation links and counts are source dependent and vary by field.	High	Use matched records, medians, agreement tests, and field stratification.
Normalized impact	Field normalization does not fully eliminate database effects.	Moderate–high	Identify the source and reference set for every normalized metric.
Metadata	No platform is uniformly most complete across all fields.	High	Calculate field-level missingness and a transparent completeness index.
ORCID	Strong for identity and interoperability; incomplete as a publication census.	High	Use ORCID for linkage and validation, not direct citation comparison.
Research assessment	Single-source rankings risk database-induced bias.	High	Triangulate sources and combine metrics with expert judgment.

Table 7. Status of the proposed hypotheses.

Hypothesis	Secondary evidence status	Primary-data decision rule
H1 Coverage differs	Supported conceptually by multiple direct comparisons; not re-tested without exports.	Cochran Q $p < .05$ plus corrected pairwise comparisons.
H2 Citation counts differ	Strongly suggested by citation-coverage studies.	Friedman $p < .05$, effect size, pairwise Wilcoxon, Bland–Altman bias.
H3 Metadata completeness differs	Supported by field-level comparisons.	Repeated-measures comparison of completeness scores and field missingness.
H4 Indicators correlate but lack perfect agreement	Consistent with comparative literature.	Spearman rho plus ICC; correlation alone cannot establish agreement.
H5 Context predicts	Plausible and repeatedly observed	Mixed-effects logistic

Hypothesis	Secondary evidence status	Primary-data decision rule
indexing	by field/type.	regression with adjusted odds ratios.
H6 DOI predicts matching	Methodologically expected; not estimated here.	Adjusted odds ratio for successful multi-source match.
H7 ORCID completeness predicts linked works	Supported in principle; requires profile-level data.	Regression of verified-linked-output proportion on profile completeness.

6. DISCUSSION

The primary conclusion is that impact divergence is database dependent. Web of Science, Scopus, and Dimensions have enough overlapping content that they can provide roughly the same information for a given context. However, the uniqueness and divergent indexing of content, left and right of the overlapping core, is large enough to shift totals, ranks, field, and assessment profiles and institutional comparisons. The data does not show a singular order of supremacy of one platform over all the others. Findings support a modular system where the coverage, curation, metadata, and identity components, as well as openness and analytics, are considered separately and impact system design differently.

Dimensions supports discovery and assists in locating applied, regional, and other non-conventional literature that may be missing from selective databases. More coverage means more data that needs to be filtered and deduplicated. Web of Science' selectivity makes a reliable analytical environment, while making some literature more difficult to locate. Scopus balances both sides as it offers advanced analytics and a wider, more, multi-disciplinary coverage. Each of their strengths and weaknesses may lead to different publication counts and different metrics for the same search query.

Citation discrepancies represent more than technical inaccuracies. Each platform builds its citation graph using indexed sources and citation references that have been properly identified and matched. Absent citing documents result in absent citations. References that are either incomplete or unmatched will result in lost citation links. When early-access and final versions are separated, citation count may be split as well. What emerges is a citation impact picture for individual platforms. While a high Spearman correlation between platforms may give reassurance to evaluators that a broad agreement will be reached, it does not eliminate the possibility of systematic distortion for platform-specific citation counts and for specific groups. That is why, for this picture to be complete, agreement analysis is required.

Findings on normalized impact are of great significance. Evaluation of citations in the context of field-normalized metrics assumes that the only thing that changes is the database, as each citation gets compared to a field and a year-specific reference set. In reality, the platforms differ in field categorization, included documents, document types, citation windows, and links. Stahlschmidt and Stephen (2022) provided an example where, even for documents available in all three repositories, completely different normalized valuations were assigned to different sectors. Hence, each normalized indicator must specify the database, the field system, the document-type specifications, and the date of the computation.

Determining the completeness of metadata establishes the feasibility of performing cross-platform assessments. Among publication keys, the DOI is the most practical, but coverage is inconsistent, and DOI strings may contain errors. Matching of titles and authors relies on different levels of robustness and is subject to various forms of disruption. The best approach balances reliability and comprehensiveness and consists of exact matching of persistent identifiers, then bibliographic fields, and then contextual information, validated by expert judgment. The sole application of any fuzzy title matching, without additional corroboration, is highly discouraged.

Reconciliation of identity is a critical service, for which ORCID is most appropriate. ORCID's design is open and researcher-centric, resulting in the persistent researcher identifier being carried across systems and organizational boundaries. Integration of ORCID-enabled systems offers the seamless linking of manuscript submissions and Crossref submissions, as well as workflows offered by funders and workflows supported by research and institutional repositories. Preliminary evidence presented by Ibrahim and Ramzy (2025) suggests that the integration of ORCID may support enhanced retrievability and visibility. An ORCID record, however, is not an up-to-date CV for the researcher and connected systems. Profile completeness, settings, duplicate records, and mandatory registration must be considered.

Bias based on discipline remains a concern. Journal-based databases flourish where journal articles dominate scholarly communication. In computer science and engineering, conference papers are more important. In the humanities and many social sciences, books and book chapters are more important. In regionally based scholarship, local-language journals are more important. In applied fields, policy reports, datasets, software, and other practice-based outputs are more important. In database comparisons across all disciplines, these biases can be hidden. The primary study should, therefore, stratify by discipline and document type. The primary study should also provide raw counts in addition to field-normalized and mission-specific counts.

The impact on universities is clear. Different sources may rely on different outputs for rankings, accreditation, annual reviews, and promotions. A single-source policy can incentivize publishing in a restricted set of journals and devalue other outputs. The universities should specify the database, the date of the search, the criteria for matching, and how missing outputs were handled. Researchers should be given an opportunity to validate their records before critical decisions are made based on these metrics.

The impact on librarians and research managers is just as important. Libraries maintain authorities that link their researchers to ORCID, Scopus, ResearcherID, and local ID systems. Research managers can implement regular profile management along with DOI clearing and maintain research systems that link and fill outputs with valid institutional data. The goal should not be to optimize a single metric, but to facilitate a more complete scholarly research record.

Responsible research assessment requires quantitative evidence to support expert judgment. DORA and the Leiden Manifesto are not anti-metrics, but are against misguided metrics and metrics without context. A multi-source model is aligned with DORA and the Leiden Manifesto because it reveals measurement uncertainty, whereas metrics without context tend to conceal measurement uncertainty. The optimal model will use ORCID for researcher's identification, at least two citation databases to capture research output, framework and career sensitive measurement normalization with explicit treatment of missing data, and qualitative assessment. When moving from a descriptive use of metrics to an evaluative use of metrics, the use of context sensitive metrics, and engagement of the community who the metrics are being used on, is of critical importance. (Pech& Delgado, 2021; Montazerian&Dorch, 2022; Rafols, 2020).

The secondary synthesis being undertaken in this research is purposely limited in nature. It does not create a new league table, nor does it generate inferential p-values from the published percentages which are inconsistent with one another. It is not accurate to pool percentages which cover different time periods, different document types, or span different academic disciplines. Instead, this research captures benchmarks as they have been reported and will develop a primary research protocol to formally test these benchmarks. This maintains the separation of evidence from proposed analysis, which is critical to the integrity of research.

6.1 Recommended Multi-Source Model

For the model being proposed, institutional records and ORCID will serve as the identity layer; Web of Science, Scopus, and Dimensions will serve as the coverage and citation layer; Crossref and OpenAlex will serve as the DOI and Open-Data validation layers, and expert review will serve as the decision layer. Metrics will be computed independently for each of the sources prior to any harmonized summary. Harmonized datasets will contain the source-specific metrics, which will be maintained in the dataset. Harmonized datasets will not contain the undocumented maximum or average.

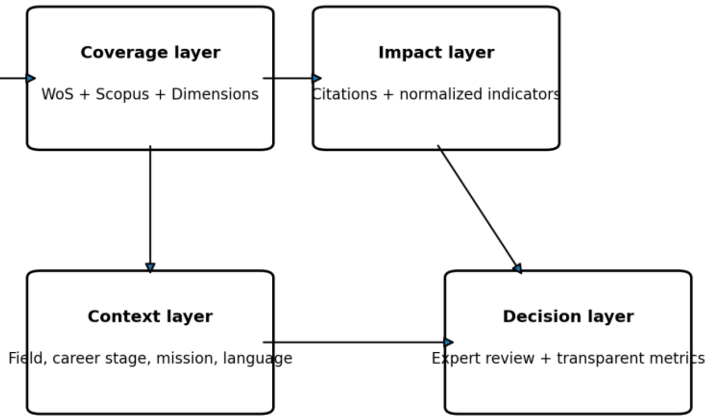

Figure 8. Recommended multi-source research-assessment model. Source: Author-developed synthesis.

Table 8. Platform-specific strengths, limitations, and appropriate uses.

Platform	Strengths	Limitations	Best-supported uses
Web of Science	Curated citation index; cited-reference tools; established field analytics	Selective coverage; subscription and export constraints; database-specific profiles	Controlled bibliometric studies, citation tracing, longitudinal evaluation with documented index set
Scopus	Broad multidisciplinary coverage; mature author/affiliation	Source-specific errors and splits; subscription	Researcher/institution profiles, multidisciplinary output and citation analysis

Platform	Strengths	Limitations	Best-supported uses
	profiles; journal and citation metrics	constraints; totals differ from WoS	
Dimensions	Broad linked data across outputs, grants and related entities; extensive discovery	Filtering and field harmonization required; access tiers affect export capability	Cross-domain research intelligence, broad recall, grant–publication linkage
ORCID	Persistent open identity; portability; interoperability; researcher control	Incomplete profiles; adoption and maintenance vary; not a citation index	Author disambiguation, trusted linkage, system integration, profile validation

7. CONCLUSION

Scopus, Web of Science, Dimensions, and ORCID each capture and characterize different constructs. The first three have bibliographic and citation infrastructures with different philosophies of indexing; the latter supplies a persistent identity and attribution infrastructure. Comparative studies conducted and verified between 2020 and 2025 suggest a considerable shared core, showing meaningful variation in breadth of sources, citation links, impact normalization, metadata, and author profiles. Dimensions tends to have greater breadth, Web of Science a more exclusive curated collection, Scopus a breadth of coverage with developed analytical profiles, and ORCID a strength in identity, linkage, and interoperability.

Concerning the research questions, the major inclusion of the shared core results in substantial citation coverage overlap, with citation visibility depending on the platform and varying by research domain. Rather than reflecting overall superiority of a single source, the degree of completeness of metadata and the profiles of researchers in different domains and subdomains are field dependent. ORCID contributes to both disambiguation and visibility of researchers, but should not be considered a citation or exhaustive publication database. Choice of database influences the metrics of researchers and institutes; therefore, a single source may be justified in cases of a focused purpose.

Justifiable assessment of a research output framework should include at least two citation databases and source-valued citations; authenticated author profiles through ORCID and institutional records; a DOI-centered, first hierarchical approach to publication matching; documented absence of data; field-specific statistics; and a blend of quantitative and qualitative data. The citation framework must be contextually bound to the purpose of the evalaution and aligned with the aims of the assessing research unit.

8. RECOMMENDATIONS

1. Develop a strategy for database selection that requires evaluative justification for each database.
2. Implement ORCID as the primary means of persistent research identity within a research information system.
3. Capture data from Scopus, Web of Science, and Dimensions within a defined temporal window and record the time of capture.
4. Keep citation counts and indicators for each source and do not replace them with maximum, average, or consolidated values.
5. Implement a DOI-first approach, supplemented by normalized data for titles, affiliations, co-authors, and source data, followed by extensive manual review.
6. Alongside total counts of publications, report overlap and unique records, as well as field-level missing data.
7. For citation data with a skewed distribution, employ medians, percentiles, rank-order (non-parametric) statistical tests, count models, and various forms of concordance (agree/disagree) measures.
8. Organize data based on discipline, document type, or citation language, as well as career stage and the institution's mission.
9. Offer the chance for researchers to challenge and/or defend their case before attribution of outputs for evaluation to inform high-impact decisions.
10. Adapt assessment methodologies to the principles of DORA and the Leiden Manifesto, as well as the principles of responsible, transparent, and verifiable assessments.

Table 9. Recommended multi-database assessment workflow.

Stage	Required action	Output
1. Governance	Define purpose, population, period, indicators, access rights, and responsible-use rules.	Approved protocol and data-management plan
2. Identity	Create researcher roster and link ORCID, local IDs, Scopus IDs, ResearcherIDs, Dimensions IDs.	Validated identity crosswalk
3. Retrieval	Export all platforms with saved queries, fields, versions, and timestamps.	Immutable raw export archive
4. Cleaning	Normalize DOI, title, names, sources, institutions, document types, dates.	Clean platform datasets
5. Matching	Apply hierarchical rules and manual review;	Master publication table

Stage	Required action	Output
	record confidence and exclusions.	
6. Analysis	Compute coverage, citations, metadata, agreement, regression, network and sensitivity analyses.	Source-specific and harmonized results
7. Validation	Researcher and expert review; inspect outliers and uncertain matches.	Validated analytical dataset
8. Reporting	Disclose source, date, limitations, missingness, effect sizes, code, and decision context.	Reproducible report

9. LIMITATIONS

The major restriction here is the lack of direct proprietary exports. For this reason, new totals of researcher-level publications, citation distributions, h-indices, network maps, regressions, and hypothesis-test p-values cannot be included in the study. The other available evidence is heterogeneous with regard to period, sample, field, and unit of analysis, thus meta-analytic pooling would not be warranted. Official coverage of the platforms is dynamic, and currently available documentation may not depict the state of the platforms in the years 2020–2025. Citation counts are mutable, and ORCID profiles are researcher-controlled. Furthermore, the subscription or API tier in question affects the fields available. The ordinal role matrix is a qualitative synthesis and should not be cited as a formal rank. These limitations are stated herein to illuminate the likely sources of illustrative material that would otherwise be mistaken for proprietary data.

10. Future Research

Future research should carry out the protocol with matched exports from multiple institutions across several countries. Citation snapshots taken longitudinally would allow for the differentiation between database growth and the accumulation of citations. Research concentrated on a specific field should include books, conference proceedings, journals in regional languages, research datasets, research software, research policy, and practice-oriented research. Additional studies should assess the impact of the trust markers and automated-update features of ORCID on profile completion, the author disambiguation accuracy on different platforms, the degree to which DOI and open metadata influence inequity in research, and the impact of research source on promotion, funding, and ranking. Where licensing allows, OpenAlex, Crossref, OpenCitations, national research information systems, institutional repositories, and citation networks should be used to add layers for validation and reproducibility.

REFERENCES

1. Aria, M., &Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics, 11(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007
2. Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. Quantitative Science Studies, 1(1), 363–376. https://doi.org/10.1162/qss_a_00018
3. Cappelli, F., et al. (2025). Deep learning approaches to author name disambiguation. International Journal on Digital Libraries. https://doi.org/10.1007/s00799-025-00428-6
4. Coalition for Advancing Research Assessment. (2022). Agreement on reforming research assessment.
5. Culbert, J. H., et al. (2025). Reference coverage analysis of OpenAlex compared to Web of Science and Scopus. Scientometrics. https://doi.org/10.1007/s11192-025-05293-3
6. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070
7. Garfield, E. (2006). The history and meaning of the journal impact factor. JAMA, 295(1), 90–93. https://doi.org/10.1001/jama.295.1.90
8. Gerasimov, I., et al. (2024). Comparison of datasets citation coverage in Google Scholar, Web of Science, Scopus, Crossref, DataCite and OpenAlex. Scientometrics. https://doi.org/10.1007/s11192-024-05073-5
9. Guerrero-Bote, V. P., Chinchilla-Rodríguez, Z., Mendoza, A., & de Moya-Anegón, F. (2021). Comparative analysis of the bibliographic data sources Dimensions and Scopus: An approach at the country and institutional levels. Frontiers in Research Metrics and Analytics, 6, 683781. https://doi.org/10.3389/frma.2021.683781
10. Hendricks, G., Tkaczyk, D., Lin, J., &Feeney, P. (2020). Crossref: The sustainable source of community-owned scholarly metadata. Quantitative Science Studies, 1(1), 414–427. https://doi.org/10.1162/qss_a_00022
11. Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., &Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. Nature, 520, 429–431. https://doi.org/10.1038/520429a

12. Houghton, F., & Foster, A. (2024). Resistance and power in Irish higher education: ORCID and the monitored university. The Journal of Academic Librarianship, 50(2), 102853. https://doi.org/10.1016/j.acalib.2024.102853
13. Ibrahim, B., &Ramzy, M. (2025). The effect of using ORCID iD on improving the visibility and retrieval of Arab University publications. Scientometrics. https://doi.org/10.1007/s11192-025-05505-w
14. Kim, J., et al. (2021). ORCID-linked labeled data for evaluating author name disambiguation at scale. Scientometrics, 126. https://doi.org/10.1007/s11192-020-03826-6
15. Kiesslich, T., Beyreis, M., Zimmermann, G., &Traweger, A. (2021). Citation inequality and the journal impact factor: Median, mean, (does it) matter? Scientometrics, 126, 1249–1269. https://doi.org/10.1007/s11192-020-03812-y
16. Kokol, P., et al. (2023). Discrepancies among Scopus and Web of Science coverage of funding information in medical journals. Journal of the Medical Library Association, 111.
17. Lim, W. M., Kumar, S., &Donthu, N. (2024). How to combine and clean bibliometric data and use bibliometric tools synergistically: Guidelines using metaverse research. Journal of Business Research, 182, 114760. https://doi.org/10.1016/j.jbusres.2024.114760
18. 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
19. Moed, H. F. (2020). Appropriate use of metrics in research assessment of autonomous academic institutions. Scholarly Assessment Reports, 2(1), 1. https://doi.org/10.29024/sar.8
20. Montazeri, A., Mohammadi, S., Hesari, P. M., Ghaemi, M., Riazi, H., &Sheikhi-Mobarakeh, Z. (2023). Preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO): A minimum requirements. Systematic Reviews, 12, 239. https://doi.org/10.1186/s13643-023-02225-6
21. Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., &Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. Profesional de la Información, 29(1), e290103. https://doi.org/10.3145/epi.2020.ene.03

22. Mukherjee, D., Lim, W. M., Kumar, S., &Donthu, N. (2022). Guidelines for advancing theory and practice through bibliometric research. Journal of Business Research, 148, 101–115. https://doi.org/10.1016/j.jbusres.2022.04.042
23. Nowakowska, M. (2025). A comprehensive approach to preprocessing data for bibliometric analysis. Scientometrics, 130, 5191–5225. https://doi.org/10.1007/s11192-025-05415-x
24. ORCID. (n.d.). About ORCID. Retrieved July 22, 2026, from https://info.orcid.org/what-is-orcid/
25. Öztürk, O. (2024). How to design bibliometric research: An overview and a framework proposal. Review of Managerial Science. https://doi.org/10.1007/s11846-024-00738-0
26. Page, M. J., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71
27. Passas, I. (2024). Bibliometric analysis: The main steps. Encyclopedia, 4(2), 1014–1025. https://doi.org/10.3390/encyclopedia4020065
28. Peroni, S., &Shotton, D. (2020). OpenCitations, an infrastructure organization for open scholarship. Quantitative Science Studies, 1(1), 428–444. https://doi.org/10.1162/qss_a_00023
29. Porter, S. R., Umbach, P. D., & Willis, C. (2025). Understanding ORCID adoption among academic researchers. Scientometrics, 130, 2783–2797. https://doi.org/10.1007/s11192-025-05300-7
30. Pranckutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. Publications, 9(1), 12. https://doi.org/10.3390/publications9010012
31. Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv. https://doi.org/10.48550/arXiv.2205.01833
32. Rethlefsen, M. L., et al. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. Systematic Reviews, 10, 39. https://doi.org/10.1186/s13643-020-01542-z
33. Scheidsteger, T., et al. (2025). How similar are field-normalized citation impact scores from OpenAlex, Web of Science, Scopus, and Dimensions? Scientometrics. https://doi.org/10.1007/s11192-025-05338-7

34. Singh, V. K., Singh, P., Karmakar, M., Leta, J., & Mayr, P. (2021). The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. Scientometrics, 126, 5113–5142. https://doi.org/10.1007/s11192-021-03948-5
35. Stahlschmidt, S., & Stephen, D. (2022). From indexation policies through citation networks to normalized citation impacts: Web of Science, Scopus, and Dimensions as varying resonance chambers. Scientometrics, 127, 2413–2431. https://doi.org/10.1007/s11192-022-04309-6
36. Teixeira da Silva, J. A. (2020). ORCID: Issues and concerns about its use for academic purposes and research integrity. Annals of Library and Information Studies, 67(4), 246–250. https://doi.org/10.56042/alis.v67i4.36267
37. Thelwall, M. (2020). The pros and cons of the use of altmetrics in research assessment. Scholarly Assessment Reports, 2(1), 2. https://doi.org/10.29024/sar.10
38. Torres-Salinas, D., Orduña-Malea, E., Delgado-Vázquez, Á., Gorraiz, J., & Arroyo-Machado, W. (2024). Foundations of narrative bibliometrics. Journal of Informetrics, 18(3), 101546. https://doi.org/10.1016/j.joi.2024.101546
39. UNESCO. (2021). UNESCO recommendation on open science. UNESCO.
40. Vieira, G. A., & Leta, J. (2024). biblioverlap: An R package for document matching across bibliographic datasets. Scientometrics, 129, 4513–4527. https://doi.org/10.1007/s11192-024-05065-5
41. Visser, M., van Eck, N. J., &Waltman, L. (2021). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. Quantitative Science Studies, 2(1), 20–41. https://doi.org/10.1162/qss_a_00112
42. Wang, T., Li, Z., Huang, S., et al. (2024). Is ORCID a reliable source for CV analysis? Exploring the data availability of ORCID academic profiles. Scientometrics, 129, 1637–1662. https://doi.org/10.1007/s11192-024-04944-1
43. Zhu, J., & Liu, W. (2020). A tale of two databases: The use of Web of Science and Scopus in academic papers. Scientometrics, 123, 321–335. https://doi.org/10.1007/s11192-020-03387-8
44. González-Alcaide, G. (2021). Bibliometric studies outside the information science and library science field: Uncontainable or uncontrollable? Scientometrics, 126, 6837–6870. https://doi.org/10.1007/s11192-021-04061-3
45. Hulland, J. (2024). Bibliometric reviews—some guidelines. Journal of the Academy of Marketing Science, 52, 935–938. https://doi.org/10.1007/s11747-024-01016-x

46. Koo, M., & Lin, S. C. (2023). An analysis of reporting practices in the top 100 cited health and medicine-related bibliometric studies from 2019 to 2021 based on proposed guidelines. Heliyon, 9, e16780. https://doi.org/10.1016/j.heliyon.2023.e16780*

47. Li, J., Deacon, C., &Keezer, M. R. (2024). The performance of bibliometric analyses in the health sciences. Current Medical Research and Opinion, 40(1), 97–101. https://doi.org/10.1080/03007995.2023.2281503

48. Montazerian, M., &Dorch, B. F. (2022). Quality and quantity in research assessment: Examining the merits of metrics. Frontiers in Research Metrics and Analytics, 7, 991550. https://doi.org/10.3389/frma.2022.991550

49. Pech, G., & Delgado, C. (2021). Screening the most highly cited papers in longitudinal bibliometric studies and systematic literature reviews: Widely used metrics versus a percentile citation-based approach. Journal of Informetrics, 15(3), 101161. https://doi.org/10.1016/j.joi.2021.101161

50. Rafols, I. (2020). S&T indicators in the wild: Contextualization and participation for responsible metrics. Research Evaluation, 29(1), 7–22. https://doi.org/10.1093/reseval/rvz032

.